S.Y.B.Sc. (C. S.)

SEMESTER - IV (CBCS)

LINEAR ALGEBRA USING PYTHON

SUBJECT CODE: USCS405

Programme Co-ordinator : **Shri Mandar Bhanushe**
Head, Faculty of Science and Technology IDOL,
Univeristy of Mumbai – 400098

Course Co-ordinator : **Ms. Mitali Vijay Shewale**
Doctoral Researcher,
Veermata Jijabai Technological Institute
HR Mahajani road, Matunga, Mumbai

Editor : **Dr. Shraddha Sable**
Assistant Professor,
IDOL, University of Mumbai- 400098

Course Writers : **Priyanka**
Assistant Professor,
J. K. College of Science and Commerce, Mumbai

: **Aarti Santosh Nayak**
Assistant Professor,
S K Somaiya college, Somaiya University,
Mumbai.

June 2023, Print - 1

Published by : Director,
Institute of Distance and Open Learning,
University of Mumbai,
Vidyanagari,Mumbai - 400 098.

DTP composed and Printed by: Mumbai University Press

CONTENTS

S.Y.B.Sc. (C. S.)

SEMESTER - IV (CBCS)

LINEAR ALGEBRA USING PYTHON

SYLLABUS

<table>
<tr><td>Course:
USCS405</td><td>TOPICS (Credits : 02 Lectures/Week: 03)
Linear Algebra using Python</td><td></td></tr>
<tr><td colspan="3">Objectives:
To offer the learner the relevant linear algebra concepts through computer science applications.
Expected Learning Outcomes:
1. Appreciate the relevance of linear algebra in the field of computer science.
2. Understand the concepts through program implementation
3. Instill a computational thinking while learning linear algebra.</td></tr>
<tr><td>Unit I</td><td>Field: Introduction to complex numbers, numbers in Python , Abstracting over fields, Playing with GF(2), Vector Space: Vectors are functions, Vector addition, Scalar-vector multiplication, Combining vector addition and scalar multiplication, Dictionary-based representations of vectors, Dot-product, Solving a triangular system of linear equations. Linear combination, Span, The geometry of sets of vectors, Vector spaces, Linear systems, homogeneous and otherwise</td><td>15L</td></tr>
<tr><td>Unit II</td><td>Matrix: Matrices as vectors, Transpose, Matrix-vector and vector-matrix multiplication in terms of linear combinations, Matrix-vector multiplication in terms of dot-products, Null space, Computing sparse matrix-vector product, Linear functions, Matrix-matrix multiplication, Inner product and outer product,</td><td>15L</td></tr>
</table>

	From function inverse to matrix inverse **Basis**: Coordinate systems, Two greedy algorithms for finding a set of generators, Minimum Spanning Forest and GF(2), Linear dependence, Basis , Unique representation, Change of basis, first look, Computational problems involving finding a basis Dimension: Dimension and rank, Direct sum, Dimension and linear functions, The annihilator	
Unit III	**Gaussian elimination**: Echelon form, Gaussian elimination over GF(2), Solving a matrix-vector equation using Gaussian elimination, Finding a basis for the null space, Factoring integers, **Inner Product**: The inner product for vectors over the reals, Orthogonality, **Orthogonalization**: Projection orthogonal to multiple vectors, Projecting orthogonal to mutually orthogonal vectors, Building an orthogonal set of generators, Orthogonal complement, Eigenvector: Modeling discrete dynamic processes, Diagonalization of the Fibonacci matrix, Eigenvalues and eigenvectors, Coordinate representation in terms of eigenvectors, The Internet worm, Existence of eigenvalues, Markov chains, Modeling a web surfer: PageRank.	**15L**

Textbook(s):

1) Coding the Matrix Linear Algebra through Applications to Computer Science Edition 1, PHILIP N. KLEIN, Newtonian Press (2013)

Additional References*:*

1) Linear Algebra and Probability for Computer Science Applications, Ernest Davis, A K Peters/CRC Press (2012).
2) Linear Algebra and Its Applications, Gilbert Strang, Cengage Learning, 4th Edition (2007).
3) Linear Algebra and Its Applications, David C Lay, Pearson Education India; 3rd Edition (2002)

1

COMPLEX NUMBER AND FIELD

Unit Structure:

1.0 OBJECTIVES

After going to this chapter, you will be able to:

- Understand the extension of real number system
- Define *i*
- Identify real and imaginary parts of a complex number
- Evaluate square root of a complex number
- Define Field.

1.1 INTRODUCTION

The concept of extension of the set of real numbers to the complex numbers was first necessitated by solution of such algebraic equations whose solutions could not be found in the set of real numbers and also to evaluate square root of a negative number.

Complex numbers were introduced by Italian mathematician Gerolamo Cardano in 1545. Leonhard Euler was first to introduce the symbol '*i*' (iota) for the square root of '-1' with the property $i^2 = -1$.

1.2 BASIC CONCEPTS OF COMPLEX NUMBER

1.2.1 Complex number: Definition and examples

Def: A number is in the form of '$a+ib$' is called a complex number, where a and b are real numbers and $i = \sqrt{-1}$.

ex. 2+3i, $\sqrt{2} + 7i, 9 - \sqrt{11}$ i

Usually a complex number is denoted by Z.

If Z = $a+ib$, then 'a' is called real part and 'b' is called imaginary part of the complex number Z and are denoted by Re(Z) and Im(Z) respectively.

A complex number whose real part is equal to 0 is called an imaginary number.

1.2.2 Algebra of complex numbers

i.) Equality of two complex numbers:

Two complex numbers $Z_1 = a_1 + ib_1$ and $Z_2 = a_2 + ib_2$ are equal iff $a_1 = a_2$ and $b_1 = b_2$.

i.e $Re(Z_1) = Re(Z_2)$ and $Im(Z_1) = Im(Z_2)$

ii.) Addition of two complex numbers:

Let $Z_1 = a_1 + ib_1$ and $Z_2 = a_2 + ib_2$ are two complex numbers. Addition of Z_1 and Z_2 is denoted as Z_1+Z_2 and defined as $Z_1 + Z_2 = (a_1+a_2) + i(b_1+b_2)$.

Example: $Z_1 = 7+2i$ and $Z_2 = 2+5i$ then $Z_1+Z_2 = (7+2i) + (2+5i) = (7+2) + i(2+5) = 9+7i$

iii.) Subtraction of two complex numbers:

Let $Z_1 = a_1+ib_1$ and $Z_2 = a_2+ib_2$ are two complex number. Subtraction of Z_1 and Z_2 is denoted as Z_1-Z_2 and is defined as $Z_1 - Z_2 = (a_1-a_2) + i(b_1-b_2)$.

Example: $Z_1 = 7+2i$ and $Z_2 = 2+5i$ then $Z_1 - Z_2 = (7+2i) - (2+5i) = (7-2) + i(2-5) = 5+(-3)i$

iv.) Multiplication of two complex numbers:

Let $Z_1 = a_1+ib_1$ and $Z_2 = a_2+ib_2$ are two complex number. Multiplication of Z_1 and Z_2 is denoted as $Z_1.Z_2$ and is defined as $Z_1 . Z_2 = .(a_1+ib_1).(a_2+ib_2)$

$= (a_1a_{2c} + ia_1b_2 + ib_1a_2 + i^2b_1b_2) = (a_1a_2 + ia_1b_2 + ib_1a_2 + (-1)b_1b_2)$ (since $i^2=-1$)

$= (a_1a_2 - b_1b_2) + (a_1b_2 + b_1a_2)i$

Multiplicative Inverse of a complex number Z = a +ib

Z^{-1} or 1/Z is called the multiplicative inverse of a non-zero complex number Z if ZZ^{-1}= 1.

$\Rightarrow Z^{-1} = \frac{1}{a+ib} = \frac{1}{a+ib} * \frac{a-ib}{a-ib} = \frac{a-ib}{a^2-b^2}$

v.) Division of two complex numbers:

Let $Z_1 = a_1+ib_1$ and $Z_2 = a_2+ib_2$ are two complex number. Division of Z_1 and Z_2 is denoted as Z_1/Z_2 and is defined as $\frac{Z_1}{Z_2} = Z_1 * \frac{1}{Z_2} = \frac{a_1a_2+b_1b_2}{{a_2}^2+{b_2}^2} + i\,\frac{a_2b_1-a_1b_2}{{a_2}^2+{b_2}^2}$

Example: Solve $\frac{1-2i}{3+4i}$.

Sol. $\frac{1-2i}{3+4i} = \frac{1+2i}{3+4i} * \frac{3-4i}{3-4i} = \frac{11-2i}{25}$

1.2.3 Conjugate, Modulus and Argument of a complex number

Let Z = a+ib is a complex number.

Conjugate: Its conjugate is denoted by $\bar{Z}$ and is defined $\bar{z}$ = a-ib.

Example: if Z= -2 +3i then $\bar{Z}$ = -2 -3i.

Modulus: The modulus(or Absolute value) of Z is denoted by |Z| and defined as $|z| = \sqrt{a^2+b^2}$

Example: Z = 5+12i,

$|z| = \sqrt{a^2+b^2} = \sqrt{5^2+12^2} = \sqrt{25+144} = \sqrt{169}$ = 13

Note: Modulus can't be negative. We always take only positive value of square root.

Argument: The argument(or Amplitude) of Z is denoted by arg(Z) or amp(Z) or 'θ' and is defined as $\tan^{-1}(\frac{b}{a})$. i.e. $\theta = \tan^{-1}(\frac{b}{a})$, when a > 0 and $\tan^{-1}(\frac{b}{a}) + \pi$, when a < 0.

Example: Z = 1 + $\sqrt{3}i$, Then amp(Z) = $\tan^{-1}(\frac{\sqrt{3}}{1}) = \frac{\pi}{3}$.

Principal argument: The principal argument of a complex number Z is Arg(Z) is equal to

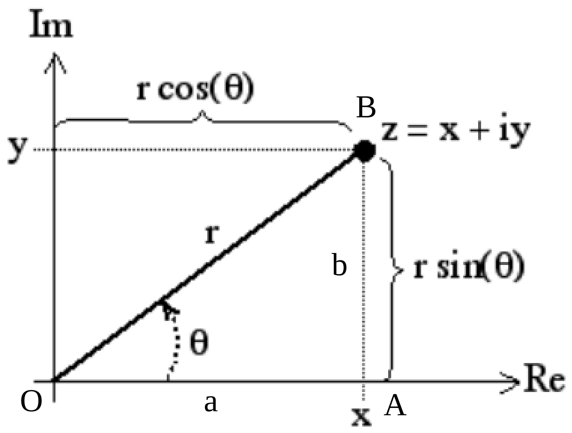
Arg(Z) = arg(Z) - 2 πn

Hence, the value of the principal argument of the complex numbers lies in the interval (-π, π).

1.2.4 Graphical representation of a complex number

A complex number Z = a+ib can be represented in a co-ordinate system known as complex plane or argand plane. We consider real part of Z (i.e.

Re(Z)=a) on X-axis (real axis) and imaginary part of Z (i.e. Im(Z)=b) on Y-axis (Imaginary axis).



From the above diagram we have OAB is a triangle.

OA= a units, AB= b units then OZ=r=$\sqrt{a^2+b^2}$ and ' θ ' is the angle between X-axis and OZ.

Cos θ = $\frac{a}{r}$, Sin θ = $\frac{b}{r}$ then tan $\theta = \frac{\sin\theta}{\cos\theta} = \frac{(b/r)}{(a/r)} = \frac{b}{a}$

Then θ =tan$^{-1}\left|\frac{b}{a}\right|$

1.2.5 Representation of a Complex number

<u>Cartesian form of a Complex number:</u>

Let Z = a+ib is a complex number. Then Z = (a, b) is the ordered pair representation or Cartesian form of complex number Z.

<u>Polar form of a complex number:</u>

Let Z = a+ib is a complex number. From the above diagram Cos θ = $\frac{a}{r}$ and Sin θ = $\frac{b}{r}$.

a = r Cosθ and b = rSin θ

Substituting these values in Z, we get

Z = a+ib= rCosθ + irSin θ

Z = r(Cos θ+i Sin θ) is called POLAR FORM of a complex number Z.

<u>Exponential Form of a complex number:</u>

Let Z = a+ib is a complex number. Then Z = r $\bullet e^{i\theta}$ is called exponential form of Z, where r is modulus of Z and θ is amplitude of Z or amp(Z).

Example: Let Z = 1+i,

Cartesian Form of Z is (1, 1)

Polar form of Z is $\sqrt{2}$(cos $\frac{\pi}{4}$ + $i$sin $\frac{\pi}{4}$), where r = $\sqrt{2}$ and $\theta = \frac{\pi}{4}$.

Exponential form of Z is $\sqrt{2}\bullet e^{i\frac{\pi}{4}}$.

1.2.6 Square root of a complex number

To find the square root of a complex number Z = a+ib, the following steps should be followed:

Step I: Let A+iB = $\sqrt{(\mathrm{a}+i\mathrm{b})}$

Step II: Squaring both sides, $(\mathrm{A}+i\mathrm{B})^2$= a+$i$b

$\Rightarrow (A^2-B^2)$ + 2AB i = a + ib

Step III: Equating real and imaginary parts from both sides;

(A^2-B^2) = a-----(i) and 2AB = b------(ii)

Step IV: Solving equations (i) and (ii), get the value of A and B.

Example: Find the square root of Z =3-4i

Soln: Given Z = $\sqrt{3-4i}$

Let $\sqrt{3-4i}$ = (a + ib) .

Squaring on both sides,

$\left(\sqrt{3-4i}\right)^2$=(a + ib)^2 .

3-4i = (a^2-b^2) + 2ab i

Comparing real and imaginary parts on both the sides;

a^2-b^2 = 3------(i) and 2ab = -4--------(ii)

$\Rightarrow$ b = $\frac{-2}{a}$

Put the value of b = $\frac{-2}{a}$ in a^2-b^2 = 3 we get, a^2- ($\frac{-2}{a}$)^2 = 3

$\Rightarrow$a^2 – ($\frac{4}{a^2}$) = 3

$\Rightarrow$a^4 – 4 = 3a^2

$\Rightarrow$a^4-3a^2 = 4

$\Rightarrow$a^4 – 4a^2 + a^2 -4 =0

$\Rightarrow$a^2(a^2-4)+1(a^2-4)=0

$\Rightarrow$ (a^2+1)(a^2-4)=0

$\Rightarrow$ (a^2+1)=0or (a^2-4)=0

$\Rightarrow$a^2 = -1 or a^2 = 4

$\Rightarrow$a = $\pm$ i(Rejected, since a must be a real number) or a=$\pm$2

if a = 2 then b= $\frac{-2}{a}$ = $\frac{-2}{2}$ = -1 and if a= -2 then b= $\frac{-2}{a}$ = $\frac{-2}{-2}$ = 1.

Therefore $\sqrt{3-4i}$ = 2-1i or -2+1i

1.3 NUMBERS IN PYTHON

In Python, there are three types of numeric.

1. Int: Int is a whole number, positive or negative, without decimals, of unlimited length.
2. Float: Float is a number, positive or negative, containing one or more decimals.
3. Complex Number: Any complex number a + *i*b is written as a + bj in python.

Variables of numeric types can be created by assigning a value to them.

Example:

```
x = 1    #int
y = 2.8  #float
Z = 2 + 1j   # complex number
```

To verify the type of any object in Python, use the type() function.

1.4 ABSTRACTING OVER FIELD

Binary Operation: A binary operation '*' is defined as a function of the product set AxA to A where for all a, b∈ A, (a*b) ∈ A.

Field: Let F is nonempty set equipped with two binary operations called addition '+' and multiplication ' ● '. Then the algebraic structure (F,+, ●) is a field if it satisfies the following postulates:

1. Closure Law: a + b ∈ F , for all a,b ∈ F
2. Associative Law: (a + b) + c = a + (b + c) , for all a, b, c ∈ F
3. Existence of identity: There exists an element e in F such that a + e = e + a = a.
4. Existence of Inverse: For each a ∈ F, there exists -a ∈ F such that a +(-a) = (-a) + a = e
5. Commutative Law: a +b = b + a for all a, b∈ F
6. Multiplication is distributive with respect to addition

   i.e. for all a, b ,c ∈ F   a●(b+c) = a ● b + a ● c (left distributive law)

   and (b+c) ● a= b ● a + b ● c (right distributive law)
7. Multiplication composition is also commutative. i.e. a●b = b●a for all a, b∈ F
8. There exists an element '1' in F such that 1● a = a = a ● 1 for all a ∈ F
9. Each none-Zero element possesses multiplicative inverse.

   Example: The set R of real numbers is a field.

1.5 PLAYING WITH GF(2):

Galois Field also known as GF(2) is the smallest field consisting only two elements **0** and **1** being the additive and multiplicative identity respectively.

The field addition in GF(2) is the logical XOR operation defined as

+	0	1
0	0	1
1	1	0

And, the field multiplication in GF(2) is the logical AND operation defined as

●	0	1
0	0	0
1	0	1

Example: 1●1+0●1+1●0+0●0+1●0 = 1 + 0 + 0 + 0 = 1

And 1●0 + 0●1 + 1●1 + 1●1 = 0 + 0 + 1 + 1 = 0

1.6 SUMMARY:

From the definition of complex number, it is clear that any imaginary number is a complex number. We can also conclude that any real number is also a complex number. In Mathematics, Complex numbers are used to find the solutions of those equations whose roots cannot be found in real number set. Algebraic operations on complex numbers are given by addition, subtraction, multiplication and division. To plot a complex number, we use complex plane that consists a coordinate system in which horizontal axis represents real component and the vertical axis represents imaginary component. The square root of a complex number is also a complex number.

1.7 REFERENCE FOR FURTHER READING:

Linear algebra and its applications, Gilbert Strang, Cengage Learning, 4th edition, 2007.

Exercise

1. If $Z_1 = 5 - 12i$ and $Z_2 = 8 + 6i$, Find the values of $Z_1 + Z_2$, $Z_1 - Z_2$, $Z_1 * Z_2$, and Z_1 / Z_2.

2. Find the conjugate, modulus and argument of the following complex numbers:

   i.) $8 - 6i$

   ii.) $5 + 12i$

   iii.) $2i$

3. Solve the following:

   i.) $(1 + 7i)(2 - 3i)$

   ii.) $(\sqrt{3} + 2i)(-2i - 1)$

   iii.) $\frac{4+3i}{2-3i}$

4. Find the square roots of the following Complex numbers:

   i.) $7-24i$

   ii.) $5+12i$

   iii.) $4-3i$

5. Solve in GF(2):

   i.) 1+1+0+1+1

   ii.) 1.1.1+0.1.1+1.1.1+0.0.0

6. Check whether the set of rational numbers and set of integers are Field or not.

❄❄❄❄❄❄❄

2

VECTORS

Unit Structure:

2.0 OBJECTIVES

After going to this chapter, you will be able to:

- Define a scalar and a vector.
- Distinguish between scalar and vector.
- Perform addition, subtraction, and multiplication by scalar on vectors.
- Represent a vector.
- Define homogeneous and non-homogeneous system of linear equations and predict nature of solution.
- Explain vector space

2.1 INTRODUCTION

A scalar is a quantity that has only magnitude. A vector is a quantity that has both magnitude and direction. We can represent a vector with a directed line segment. The arrow indicates the direction and the length is the magnitude of the vector.

2.2 VECTORS ARE FUNCTIONS

Vectors can be represented as a function. It is called a vector function. The domain of the vector function consists of one or more variables and returns a vector. A vector function of a single variable in R^2 and R^3 have the form, r(t) = <f(t), g(t)> and r(t) = <f(t), g(t), h(t)> respectively, where f(t), g(t) and h(t) are called the component functions. In general, a vector function of single variable in R^n has the form:

$v(t) = \langle f_1(t), f_2(t), f_3(t), \ldots, f_n(t)\rangle$ where $f_1(t), f_2(t), f_3(t), \ldots, f_n(t)$ are n-components.

The domain of a vector function is the subset of real numbers and set of all t's for which all the component functions are defined. The range is a vector.

2.3 VECTOR ADDITION AND SCALAR MULTIPLICATION

2.3.1 Vector Addition:

Vector addition is the operation of adding two or more vectors together. In Linear Algebra, vectors are given in their components form. Vector addition can be performed simply by adding the corresponding components of the vectors, so in R^n, if U and V are two vectors with n-components $U = (u_1,u_2,\ldots,u_n)$ and $V = (v_1,v_2,\ldots,v_n)$ then,

$U + V = (u_1+ v_1, u_2 + v_2,\ldots,u_n + v_n)$.

Vector addition is possible if both the vectors have same number of components.

2.3.2 Scalar-vector Multiplication:

When a vector V is multiplied by a scalar quantity k, its magnitude becomes k-times of the original vector but the direction depends on the sign of k. If k is positive, then kV has the same direction of V, but if k is negative, kV has the opposite direction of V. In linear Algebra, to multiply a vector V having components $(v_1,v_2,\ldots,v_n)$ by a scalar k means to multiply each component of the given vector by the scalar k.

$\Rightarrow kV = k(v_1,v_2,\ldots,v_n) = (kv_1, kv_2,\ldots, kv_n)$

2.3.3 Combining vector addition and scalar multiplication:

Vector addition and scalar multiplication simultaneously can be performed by following these steps:

1. Complete the scalar multiplication first by multiplying each component of the vector V by the scalar k.
2. Then, perform the vector addition by adding corresponding components of vectors that have been found after completing step 1.

   Example: If u = (2, 3, -1) and v = (6, -3, -2), then find

   (a.) (u + v)(b.) 2u + 3v (c.)(u – v)

   Solution: (a.) (u + v) = (2, 3, -1) + (6, -3, -2) = (2 + 6, 3 + (-3), (-1) + (-2)) = (8, 0, -3)

   (b.) 2u + 3v = 2 (2, 3, -1) + 3(6, -3, -2) = (4, 6, -2) + (18, -9, -6)

   = (4 + 18, 6 +(-9), (-2) + (-6)) = (22, -3, -8)

   (c.) (u – v) = (2, 3, -1) - (6, -3, -2) = (2 - 6, 3 - (-3), (-1) - (-2)) = (-4, 6, 1)

2.4 DICTIONARY BASED REPRESENTATION OF VECTORS

A vector is a function from some domain D to a field. In Python, it can be represented by a dictionary. For this, define a Python class Vec with two variables f (the function represented by Python dictionary) and D(the domain of the function represented by a python set.

```
class Vec:
    def__init__(self, labels, function):
        self.D = labels
        self.f = function
```

can create

```
>>> Vec({'A', 'B', 'C'}, {'A': 1})
```

Can assign an instance to a variable and subsequently access the two fields of v,

```
>>> v = Vec({'A', 'B', 'C'}, {'A': 1})
>>> for d in v.D:
…   if d in v.f:
…       print(v.f[d])
…
```

2.5 DOT PRODUCT

The dot product of two vectors with n-components U = (u_1,u_2,……..,u_n) and V = (v_1,v_2,……..,v_n) is denoted as U.V and defined as U.V = (u_1.v_1 + u_2.v_2 + ……..+ u_n.v_n).

U.V is a scalar quantity and it follows commutative law of multiplication. That is, U.V = V.U

Example 1; Find dot product of (1, 2) and (3, 4).

Solution: Let U = (1, 2), V = (3, 4), then U.V = (1, 2)•(3, 4) = (1*3 + 2*4) = 11

Example 2: The dot product of two vectors from R^3 where u = (1, -1, 2) and v = (2, -3, 4).

Solution: u . v = (1, -1, 2)•(2, -3, 4) = 1* 2 + (-1)*(-3) + 2*4 = 2 + 3 + 8 = 13

Example 3: Let u =11001 and v=10110 are two vectors over GF(2), find their dot product.

Solution: u.v = (11001)•(10110) = (1*1 + 1*0 + 0*1 + 0*1 + 1*0) = (1 + 0 + 0 + 0 + 0) = 1.

2.6 SOLVING A TRIANGULAR SYSTEM OF LINEAR EQUATIONS

Consider the system of n linear equations:

$a_{11}x_1 + a_{12}x_2 + \ldots\ldots\ldots + a_{1n}x_n = K_1$------------------------(i)

$a_{21}x_1 + a_{22}x_2 + \ldots\ldots\ldots + a_{2n}x_n = K_2$------------------------(ii)

$a_{n1}x_1 + a_{n2}x_2 + \ldots\ldots\ldots + a_{nn}x_n = K_n$------------------------(n^{th})

Containing the n unknowns x_1, x_2…, x_n. It is called a linear system of equations. The leading unknown in all equations is x_1 and the leading co-efficient of equations are a_1, a_2, …, a_n respectively.

2.6.1 Lower Triangular System: The linear system of equations is called lower triangular system of equations if leading unknown in all equations is x_1 and the leading co-efficient of equation (i) is a_{11}, leading co-efficient of equation(ii) is a_{21} and so on i.e the general form of triangular system of n linear equation having n unknown is

$a_{11}\,x_1 = K_1$

$a_{21}\,x_1 + a_{22}x_2 = K_2$

.

.

.

$a_{n1}\,x_1 + a_{n2}\,x_2 + \ldots\ldots\ldots + a_{nn}\,x_n = K_n$

The lower triangular system of equations can be solved by forward substitution method i.e First we have to calculate value of x_1 by 1st equation.

$\Rightarrow a_{11}\, x_n = K_1$

$\Rightarrow x_1 = \dfrac{k_1}{a_{11}}$

Then the value of x_2 is obtained by putting the value of x_1 in 2nd equation and then solving it.

So, we proceed up to last equation where we can get value of x_n by substituting the values of $x_1, x_2, \ldots\ldots\ldots x_{n-1}$.

<u>2.6.2 Upper Triangular System:</u> The linear system of equations is called upper triangular system of equations if leading unknown in the first equation is x_1 , leading unknown in the second equation is x_2, that of the third equation isx_3,……..and so on .And the leading co-efficient of equation (i) is a_{11}, leading co-efficient of equation(ii) is a_{22} and so on i.e the general form of triangular system of n linear equation having n unknown is

$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots\ldots + a_{1n}x_n = K_1$

$a_{22}x_2 + a_{23}x_3 + \ldots\ldots + a_{2n}x_n = K_2$

$a_{33}x_3 + \ldots\ldots + a_{3n}x_n = K_3$

$a_{nn}\, x_n = K_n$

The upper triangular system of equations can be solved by backward substitution method i.e First we have to calculate value of x_n by n^{th} equation.

$\Rightarrow a_{nn}\, x_n = k_n$

$\Rightarrow x_n = \dfrac{k_n}{a_{nn}}$

Then the value of x_{n-1} is obtained by putting the value of x_n in 2nd last equation and then solving it.

So, we proceed up to first equation where we can get value of x_1 by substituting the values of $x_2, x_3, \ldots\ldots\ldots x_{n-1,}\, x_n$.

Example 1: $5\,x_1 = 15$, $4x_1 + 2x_2 = 10$, $3x_1 + 5x_2 + 2x_3 = 18$

Solution: The given system is lower triangular system of linear equations having 3 unknowns. Hence by forward substitution method;

$5\,x_1 = 15 \Rightarrow x_1 = \frac{15}{5} = 3$

By substituting value of x_1 in equation (ii), $4*3 + 2x_2 = 10 \Rightarrow x_2 = -1$

Now replacing values of x_1 and x_2 in equation (iii), $3*3 + 5*(-1) + 2x_3 = 18 \Rightarrow x_3 = 7$

Example 2: $x_1 + 2x_2 + x_3 = 8$

$$3x_2 + 4x_3 = 18$$

$$7x_3 = 21$$

Solution: The given system is upper triangular system of linear equations having 3 unknowns. Hence by backward substitution method;

$7x_3 = 21 \Rightarrow x_3 = 3$

Substitute the value of x_3 in second equation we get,

$3x_2 + 4x_3 = 18 \Rightarrow x_2 = 2$

Substituting x_2 and x_3 in first equation we get,

$x_1 + 2x_2 + x_3 = 8 \Rightarrow x_1 = 1$

2.7 LINEAR COMBINATION

Let $v_1, v_2, \ldots, v_n$ are n vectors, then the combination $(\propto_1 v_1 + \propto_2 v_2 + \ldots + \propto_n v_n)$ is called a linear combination of the vectors $v_1, v_2, \ldots, v_n$ where $\propto_1, \propto_2, \ldots, \propto_n \in F$.

It can be geometrically interpreted as the vectors $v_1, v_2, \ldots, v_n$ will be added with each other after scaling by $\propto_1, \propto_2, \ldots, \propto_n$ times respectively.

Example 1: Express W = (6, -2, 5) as a linear combination of $v_1 = (-2, 1, 3)$ and $v_2 = (3, 1, -1)$ and $v_3 = (-1, -2, 1)$.

Solution: $(6, -2, 5) = a_1(-2, 1, 3) + a_2(3, 1, -1) + a_3(-1, -2, 1)$

$$= (-2a_1, a_1, 3a_1) + (3a_2, a_2, -a_2) + (-a_3, -2a_3, a_3)$$

$$= -2a_1 + 3a_2 - a_3,\ a_1 + a_2 - 2a_3,\ 3a_1 - a_2 + a_3$$

Comparing respective components of both sides, we get

$$-2a_1 + 3a_2 - a_3 = 6 \text{ ---------(i)}$$

$$a_1 + a_2 - 2a_3 = -2 \text{ -----------(ii)}$$

$$3a_1 - a_2 + a_3 = 5 \text{ ------------(iii)}$$

Solving these equations by using Cramer's rule or matrix method, we get

$a_1 = 9/5$, $a_2 = 23/5$ and $a_3 = 21/5$

$\Rightarrow (6, -2, 5) = 9/5(-2, 1, 3) + 23/5(3, 1, -1) + 21/5(-1, 2, 1)$.

Example 2: Express W=(4, 3) as a linear combination of v_1=(2, 3) and $v_2 = (0, 1)$.

Solution: (4, 3) = a_1(2, 3) + a_2 (0, 1)

⇒ (4, 3) = ($2a_1$, $3a_1$) + (0, a_2)

⇒ (4, 3) = ($2a_1$, $3a_1 + a_2$)

Comparing respective components of both sides, we get

$2a_1 = 4 \Rightarrow a_1 = 2$ and $3a_1 + a_2 = 3 \Rightarrow -3$

⇒ (4, 3) = 2(2, 3) + (-3)(0, 1)

2.8 SPAN

The set of all linear combinations of finite sets of elements of S is called Linear Span of S and is denoted by L(S) or [S]

L(S) = { $\alpha_1 v_1 + \alpha_2 v_2 + \ldots\ldots\ldots\ldots + \alpha_n v_n : v_1, v_2, \ldots\ldots v_n \in S \;\; \alpha_1, \alpha_2, \ldots\ldots\ldots\ldots \alpha_n \in F$}

Example 1: Find the Span of a subset S= {(1, 0, 0), (0, 1, 1)} of vector space V_3.

Solution: L(S)= **{** α_1 (1,0,0) + α_2 (0,1,1)}

= { $\alpha_{1,}0, 0) + (0, \alpha_2, \alpha_2)$}

={ ($\alpha_1, \alpha_2, \alpha_2$)}

⇒ The linear span of the given subset of v_3 is the element of xyz-plane, whose y and z co-ordinates are same.

Example 2: Find the span of subset S={ (1, 3), (0, 2) } of vector space V_2 show that (2,8) belongs to span S.

Solution: L(S) = { $\alpha_1 (1, 3) + \alpha_2 (0, 2)$}

= **{ (**α_1 , 3 α_1) + (0, $2\alpha_2$)} = { (α_1 , $3\alpha_1 + 2\alpha_2$)}

If (2, 8) ∈ L(S) then α_1 =2 and $3\alpha_1 + 2\alpha_2 = 8 \Rightarrow \alpha_2 = 1$

⇒(2, 8) = 2(1, 3) + 1(0, 2)

Example 3: Show that the subset S = {(1, 0, 0), (0, 1, 0), (0, 0, 1)} of V_3 spans the entire vector space V_3.

Solution: Let (a, b, c) ∈ V then (a, b, c) = a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1)

Thus (a, b, c) ∈ L(S).

Hence the subset S span the entire vector space.

Example 4: $v_1 = (1, 0, 1)$, $v_2 = (2, 1, 4)$, $v_3 = (1, 1, 3)$ do not span vector space.

Solution: Let (a, b, c) $\in$ V and (α_1 , α_2 , α_3) $\in$ F.

And S = {(1, 0, 1), (2, 1, 4), (1, 1, 3)}

(a, b, c) = α_1 $(1, 0, 1)$ + α_2 $(2, 1, 4)$ + $\alpha_3(1, 1, 3)$

(a, b, c) = (α_1, 0 , α_1) + ($2\alpha_{2,}$, α_2 $4\alpha_2$) + (α_3, α_3 , 3 α_3)

(a, b, c)= ($\alpha_1 + 2\alpha_2 + \alpha_3$, α_2+α_3 , $\alpha_1 + 4\alpha_2 + 3\alpha_3$)

a = $\alpha_1 + 2\alpha_2 + \alpha_3$ –(i)

b = α_2+α_3 ---(ii)

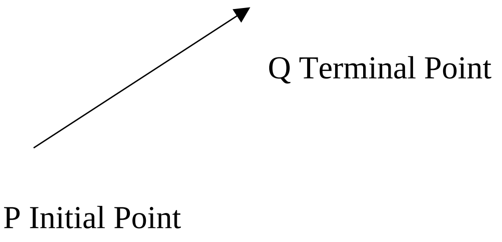
c = $\alpha_1 + 4\alpha_2 + 3\alpha_3$---(iii)

Solving (i) and (iii), α_2+α_3 = $\frac{c-a}{2}$ $\Rightarrow$ $\frac{c-a}{2} = b$ $\Rightarrow$ c-a=2b

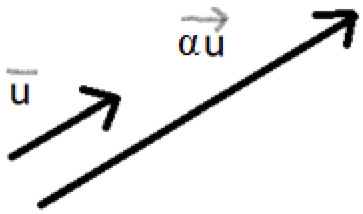
$\Rightarrow$ the set S does not span the entire v_3, but it spans a subset of V whose co-ordinates (a, b, c) satisfy the relation a + 2b – c = 0.

2.9 GEOMETRY OF SET OF VECTORS

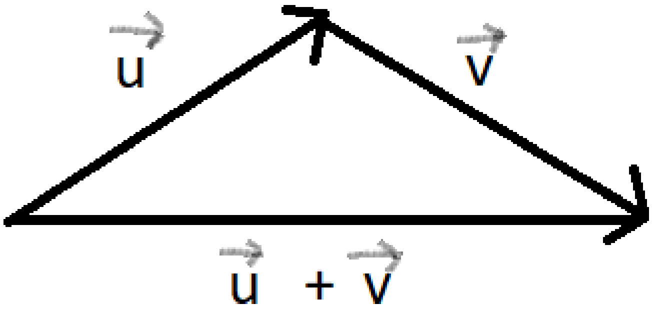
In geometry, vectors are represented by an arrow. The head of the arrow indicates its direction and length describes the magnitude of the vector.



If we multiply a vector u by a scalar $\propto$, then the length of the vector stretches by the factor $\propto$. If $\propto$ is negative, then the direction of the vector will be reversed.



If the vector u is added to vector v, then their sum is the new vector (u + v) that paints from the tail of u to the tip of v as shown:



The length or magnitude of an n-vector is defined as $||v|| = \sqrt{v.v}$

i.e if $v = (v_1, v_2, \ldots, v_n)$, then $||v|| = \sqrt{v_1^2 + v_2^2 + v_n^2} = \sqrt{\sum_{i=1}^{n} v_i^2}$

The angle θ between two n-vectors is determined by $u . v = ||u||\ ||v||. \cos\theta$

2.10 VECTOR SPACE

Binary Composition: Binary composition is an operation of two elements of the set whose domains and co-domain are in the same set.

The composition '*' is called internal composition if $a*b \in A, \forall a, b \in A$ and $a*b$ is unique.

The composition 'o' is called external If $a \text{ o } \alpha \in V$, for all $a \in F$ and for all $\alpha \in V$ and $a \text{ o } \alpha$ is unique.

<u>Vector Space:</u> Let V is a non-empty set equipped with two binary operations '**.**' (external composition) defined as scalar multiplication and '+' (internal composition) defined as addition of vectors. Then V is called a Vector space over a field F if it satisfies the following postulates:

i.) Closure Law: $(\propto + \beta) \in V: for\ all\ \propto, \beta \epsilon V$

ii.) Associative Law: $(\propto + \beta) + \gamma = \propto + (\beta + \gamma)$

iii.) Existence of Identity: There exists an element $e \epsilon V$ such that $\propto + e = e + \propto = \propto$.

iv.) Existence of Inverse: For each element $\propto \in V$, there exist an element β such that

$(\propto + \beta) = (\beta + \propto) = e$

v.) Commutative Law: $(\propto + \beta) = (\beta + \propto)$

vi.) (Closure law with respect to scalar multiplication): $a \propto \in V$ for all $a \in F$ and for all $\propto \in V$

vii.) $a(\propto + \beta) = a \propto + a\beta$, for all $a \in F$ and for all $\propto, \beta \in V$

viii.) $(a + b) \propto = a \propto + b \propto$, for all $a \in F$ and for all $\propto \in V$

ix.) $(ab)\propto = a\,(b \propto)$, for all $a, b \in F$ and for all $\propto \in V$

x.) $1. \propto = \propto$, for all $\propto \epsilon V$ And 1 is the unity element of the field F.

Example1: The set of complex numbers 'C' is a vector space over the field of real numbers R.

Solution: Let X = a+ib ∈ C, Y = c+id ∈ C, Z = p+iq, where a, b, c, d, p, q ∈ R.

i) Closure law: (X+Y) = (a+ib) + (c+id)=(a+c)+i(b+d) ∈C

ii) Associative law: X+(Y+Z)=(X+Y)+Z

L.H.S. = (a+ib) + ((c+id)+(p+iq)) = (a+ib)+((c+p)+i(d+q))

= (a+p+c) + i(b+d+q) = ((a+c)+p) + i ((b+d)+q)

= ((a+c)+i(b+d)) + (p+iq) = ((a+ib) + (c+id)) + (p+iq) = (X+Y)+Z = R.H.S.

iii) Existence of Identity: let X = a+ib ∈ C, ∃ an element e = 0+0i ∈ C such that

X + e = e+ X =X

(a + ib) + (0 + 0i) = (0 + 0i) + (a + ib) = (a + ib)

(a + 0) + i(b + 0) = (0 + a) + i(0+b) = (a + ib)

iv) Existence of Inverse: Let X=(a+ib) ∈ C, ∃ an element X′ = - (a + ib) ∈ C

Such that X+X'=X'+X=e(where e =0+0i)

(a + ib) + [-(a + ib)] = (-a + a) + i(-b + b) = 0+0i = e

v) Commutative law: Let X=(a+ib) ∈ **C ,** Y = (c+id) ∈ **C** where a, b, c, d ∈ R.

Consider X+Y=(a+ib)+(c+id) = (a+c)+i(b+d)

= (c+a)+i(d+b) = (c+id)+(a+ib) = Y+X

vi) Closure law w.r.t. scaler multiplication under vector addition: Let ∀ K ∈ R, ∀ X ∈ C**,** such that KX ∈ **C,** where K is any scaler value.

Consider K X = K(a+ib) = (ka + $i$kb)

= $(a_1 + ib_1)$ ∈ **C**

vii) Closure law w.r.to scalar multiplication under vector addition:

K (X+Y) = K ((a+ib)+(c+id)) = K (a+ib) + K(c+id) =K X + KY

viii.) Let ∀ k_1, k_2 ∈ R, X = (a+ib) ∈ C, Such that

(k_1+k_2)X=$(k_1 + k_2)$(a+ib) = k_1(a+ib) +k_2(a+ib) = k_1X +k_2X.

ix.) Let $\forall\, k_1, k_2 \in R$, X = (a+ib) ∈ C, Such that $(k_1.k_2)$ X = $(k_1.k_2)$(a+ib) = $k_1.(k_2$(a+ib) =$k_1.(k_2(X))$

x.) Multiplication with unity: ∀ X ∈ C, ∃1 ∈ R is the unity element such that 1.X=1.(a+ib) = (a+ib) = X

Since, the set of complex numbers satisfies all postulates. Hence, the set of complex number 'C' is a vector space over the field of real number R.

Example 2: Check whether the set of all pairs of real numbers of the form (1, x) with operation (1, y) + $(1, y') = (1, y + y')$ and k(1, y)=(1, ky) is a vector space.

Solution: Let (1,x) , $(1, x') \in R^2$

i.) Closure Property: Consider $(1, x_1) + (1, x_2) = (1, x_1 + x_2)$

$= (1, x_1 + x_2) \in R^2$ as $(x_1 + x_2) \in R^2$

ii.) Associative Property: Set of real numbers satisfies Associative Property.

iii.) Existence of Identity: $\exists(1, 0) \in R^2$ and $\forall(1, x) \in R^2$ such that

(1, 0) + (1, x) = (1, x) + (1, 0) = (1, x)

iv.) Existence of Inverse: $\exists(1,-x) \in R^2$, $\forall(1,x) \in R^2$ such that

(1, x) + (1, -x) = (1, -x) + (1, x) = (1, 0)

v.) Commutative Property: (1, x) , $(1, x') \in R^2$ such that $(1,x) + (1, x')=(1, x+ x')$

$= (1, x+ x') = (1, x'+x) = (1, x') + (1, x)$

Hence commutative Property is satisfied

vi.) Closure law w.r.t. scalar multiplication: k(1, y)=(1, ky), by the definition .

vii.) Closure law w.r.to scalar multiplication under vector addition:

$a[(1, x) + (1, x')] = a[1, x+x'] = [1, a(x+x')] \in R^2, \forall a \in R$

(by the definition of addition)

viii.) (a + b)●(1, x) = [1, (a+b)x] (by the definition)

and $[1, (a + b)x] \in R^2$, $\forall$ a, b $\in R^2$

ix.) (a●b) [1, x] = [1, (a●b)x] (by the definition)

= a(1, bx) = a (b[1, x])

x.) Multiplication with unity:

1●[1, x] = [1, 1●x] = [1, x] where 1∈R

Since all the postulates for becoming the vector space satisfied and hence it is a vector space.

2.11 LINEAR SYSTEMS-HOMOGENEOUS AND OTHERWISE

Linear algebra is a systematic study of the theory and applications of linear system of equations. Consider the system of m linear equations

$$a_{11}\,x_1 + a_{12}\,x_2 + ------ + a_{1n}\,x_n = b_1$$

$$a_{21}\,x_1 + a_{22}\,x_2 + ------ + a_{2n}\,x_n = b_2$$

$$a_{m1}\,x_1 + a_{m2}\,x_2 + ------ + a_{mn}\,x_n = b_m$$

having n unknowns x_1 , x_2 , …… , x_n . To determine whether the system has a solution or not, we check the ranks of the matrices,

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots\ldots. & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots. & a_{2n} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ a_{m1} & a_{m2} & \ldots\ldots. & a_{mn} \end{pmatrix}$$

And

$$B = \begin{pmatrix} a_{11} & a_{12} & \ldots\ldots. & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots\ldots. & a_{2n} & b_2 \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots & \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots & \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots & \\ a_{m1} & a_{m2} & \ldots\ldots. & a_{mn} & b_m \end{pmatrix}$$

Where A is the coefficient matrix and B is the augmented matrix of the system of equations.

Procedure to test the consistency of equations in n unknowns:

Let the rank of A be r and rank of B be r'.

1.) If r $\neq$ r' , there is no solution of the system of equations. This implies that equations are inconsistent.

2.) If r = r' = n (number of unknowns), there is a unique solution. This implies that equations are consistent.

3.) If r = r' < n, there is infinite number of solutions. This implies that equations are consistent.

<u>System of linear homogeneous equations</u>:

Consider the homogeneous linear equations

$$a_{11}\, x_1 \;+\; a_{12}\, x_2 + \text{------} + \; a_{1n}\, x_n \;=\; 0$$

$$a_{21}\, x_1 \;+\; a_{22}\, x_2 + \text{------} + \; a_{2n}\, x_n \;= 0$$

$$a_{m1}\, x_1 \;+\; a_{m2}\, x_2 + \text{------} + \; a_{mn}\, x_n \;= 0$$

To know the nature of the solutions of equation (ii), we check the rank of coefficient matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots\ldots. & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots. & a_{2n} \\ & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & & \\ & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & & \\ & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & & \\ a_{m1} & a_{m2} & \ldots\ldots. & a_{mn} \end{pmatrix}$$

Let rank (A) = r.

1.) If r = n, the equations (ii) have only trivial zero solution. This implies that

$$x_1 = \; x_2 = \text{----} = x_n \;\; = 0$$

2.) If r < n, the equations (ii) have infinite number of solutions.

We can conclude that for a homogeneous system of equations, if det (A) ≠ 0, there exists only a trivial zero solution otherwise infinitely many solutions will exist.

Example 1 : Consider the following system of equations and Find the nature of solution without solving it.

i.) $x_1 + x_2 = 6$ and $2x_1 + 2x_2 = 12$

ii.) $x_1 + x_2 = 5$ and $x_1 - x_2 = 1$

Solution: i.) The system of equations can be written in matrix form as

$$\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 12 \end{pmatrix}$$

Coefficient matrix A = $\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ and Augmented matrix B = $\begin{pmatrix} 1 & 1 & 4 \\ 2 & 2 & 8 \end{pmatrix}$

Here det A = 0, rank A = 1 and rank B = 1, So $r = r' < n$ (number of variables)

Hence there is infinite number of solutions for this system.

ii.) Here A = $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ and B = $\begin{pmatrix} 1 & 1 & 5 \\ 1 & -1 & 1 \end{pmatrix}$

Since rank A = rank B = n(number of variables), $r = r' = n$

Hence there exists a unique solution of the system.

2.12 SUMMARY

In a very simple definition, vector can be assumed as an arrow that points in space. A vector that contains n elements is called n-vector. Vector addition satisfies algebraic properties like commutative and associativity. Scalar-vector multiplication stretches the direction of a vector and this process is called scaling. These properties of vectors give the data analyst a nice way to conceptualize many list of numbers in a visual way to be clear about patterns in data.

2.13 REFERENCE FOR FURTHER READING

Linear algebra and its applications, Gilbert Strang, Cengage Learning, 4th edition, 2007.

Exercise

Q.1 For the given pairs of vectors ,find vector u + v, u – v , v – u ,2u + 3v , -2u – 7v

(i) u = (2, 8) and v = (3, 1) (ii) u = (-1, 3) and v = (8, -2)

(iii) u = (-3, 4) and v = (1, -2) (iv) u = (2, -9) and v = (-8, 1)

Q.2 For each of the following pairs of vectors u and v, Evaluate their dot product u. v.

(i)u = (2, 5) and v = (4, -1) (ii) u = (1, 2, -1) and v = (1,-1, 0)

Q.3 Solve the following triangular system of linear equation :

(i) x_1-3x_2-2x_3 = 15 (ii) 2x_1-3x_2+5x_3-2x_4 =9

2x_2+4x_3 = 8 5x_2+x_3-3x_4 =9

10x_3 = 30 7x_3-x_4 =9

2x_4 =8

Q.4 Determine whether the following set of vectors span vector space R^3

(i)v_1(2, 2, 2) , v_2(0, 0, 3) , v_3(0, 1, 1)

(ii)v_1(1,0,0) , v_2(0,1,0) , v_3(1,1,0)

Q.5 Check whether the following sets are vector space or not:

i.) {(x, y, z):x, y, z ∈ R, x + y + z = 0}

ii.) All mxn matrices whose entries are real.

❄❄❄❄❄❄❄

3

MATRIX

Unit Structure:

3.0 OBJECTIVES

After going through this chapter, students will able to learn

- To understand what are matrices
- To deal with various types of matrices using vectors
- To learn various concepts and applications of matrices using python

3.1 INTRODUCTION

This unit will take thorough out the concepts of matrices – some traditional while some are new in terms of vectors , various operations and other concepts.

3.2 MATRICES

In this section definition of matrix will be reviewed and a new notation in terms of python list will be introduced.

3.2.1 Definition

Traditionally matrices means some set of rows and columns with various entries like real numbers, complex number etc.

For example :

$$\begin{bmatrix} 1 & 0 & 1 \\ 2 & 4 & 5 \\ -2 & 3 & 1 \end{bmatrix} \text{ or } \begin{bmatrix} 1+i & -3 \\ 2+2i & 3-1 \end{bmatrix}$$

The first matrix is called as a 3x3 matrix over field F

In first example above there are 3 rows and 3 columns. First row or Row 1 is [1 0 1] , similarly column 1 is $\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$ and so on.

In general, a matrix with m rows and n columns is called mxn matrix. For a i,jth element is defined to the element in ith row and jth column . Traditionally if matrix is given by A, this element is written as A_{ij}.

Instead Python notation will be used throughout A[i,j].

So ,Row vector i will be : [A[i, 0], A[i, 1], A[i, 2], · · ·, A[i,m − 1]]

and column vector j will be : [A[0, j], A[1, j], A[2, j], · · ·, A[n − 1, j]]

For example : if we consider same matrix $\begin{bmatrix} 1 & 0 & 1 \\ 2 & 4 & 5 \\ -2 & 3 & 1 \end{bmatrix}$ then

Row vector 1 will be : [[1,0,1]] and column vector 1 will be [[1,2,-2]]

Entire matrix can be represented as list of lists as :

[[1,0,1], [2,4,5], [-2,3,1]]

In general a matrix can be represented as list L :

A[i, j] = L[i][j] for every $0 \le i < m$ and $0 \le j < n$

3.2.2 Column Space and Row Space

Matrices can be viewed from various angles like pack of rows or pack of columns etc. There are two ways of interpreting a matrix in terms of vector space. Similarly, there are two vector spaces associated with any given matrix:

Definition : For any matrix A :

1. Column space of A, written Col A, is the vector space spanned by the columns of M,
2. Row space of A, written Row A, is the vector space spanned by the rows of M.

For example : if we consider same matrix $\begin{bmatrix} 1 & 0 & 1 \\ 2 & 4 & 5 \\ -2 & 3 & 1 \end{bmatrix}$ then

Col A will be span of [[1,2,-2], [0,4,3], [1,5,1]]

And Row A will be span of [[1,0,1],[2,4,5], [-2,3,1]]

3.2.3 Transpose

Transpose of a matrix means interchanging its rows and columns.

Definition : The transpose of a matrix A, denoted by A^T is defined by

$(A^T)i,j = Aj,i$ for every i ,j .

For example : transpose of matrix $\begin{bmatrix} 1 & 0 & 1 \\ 2 & 4 & 5 \\ -2 & 3 & 1 \end{bmatrix}$ is $\begin{bmatrix} 1 & 2 & -2 \\ 0 & 4 & 3 \\ 1 & 5 & 1 \end{bmatrix}$

3.2.4 Vectors

Matrices can be represented as vectors . If AxB is a matrix over the field F then it can be represented as vector over F. Later it can be used to perform vector operations like addition of vectors, multiplication of scalar- vector.

For example : if we consider matrices A = $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 5 \end{bmatrix}$ and B = $\begin{bmatrix} 0 & 2 & 1 \\ 5 & -1 & 2 \end{bmatrix}$ then A + B = $\begin{bmatrix} 1 & 4 & 2 \\ 7 & 0 & 7 \end{bmatrix}$ i.e corresponding elements get added.

Note matrices should have same dimensions i.e number of rows and columns.

Similarly, scalar matrix multiplication is :

A = $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 5 \end{bmatrix}$ and scalar α = 3 then αA = $\begin{bmatrix} 3 & 6 & 3 \\ 6 & 3 & 15 \end{bmatrix}$

3.3 MULTIPLICATION IN TERMS OF VECTORS

In this section the concept of matrix multiplication by vectors will be discussed. There are two ways in which this can be done :

- Matrix-Vector multiplication i.e multiply a matrix by vector.
- Vector- Matrix multiplication i.e multiply a vector by matrix.

In the following section both these concepts will be discussed with two definitions for each : one in terms of dot products and another in terms linear combinations; both of which are equivalent.

3.3.1 Matrix-Vector Multiplication

Definition : In terms of Linear Combination :

Let M be RxC matrix over field F. Let ν be a vector of dimension C. Then M ∗ ν is the linear combination $\sum_{c \in C} v\,[c](column\ c\ of\ M)$

<u>Note :</u>

1) If M is R × C matrix but ν is not of dimension C i.e it is not a C-vector then the product M ∗ ν is illegal.

2) In the case of traditional-matrix, if M is m × n matrix over F then M ∗ ν is legal only if ν is n-vector over F i.e the number of columns of the matrix and the number of entries of the vector must be same.

Example 1 : Suppose M = $\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \end{bmatrix}$ and ν = $[1, \quad -1, \quad 0]$

Then M ∗ ν can be computed since M is 2x3 and ν is 3x1 and result is :

M ∗ ν = $\sum_{c \in C} v\,[c](column\ c\ of\ M)$

= 1[1,2] + (-1) [0,1] + 0[1,3] = [1,2] – [0,1] + [0,0]

= [1,1]

Example 2 : Suppose M = $\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \end{bmatrix}$ and ν = $[1, \quad 0]$

Then M ∗ ν cannot be computed since M is 2x3 and ν is 2x1 and result is not valid (by note 1)

Definition : In terms of Dot Product:

Let M be RxC matrix over field F. Let u be a vector of dimension C. Then M∗u is the R-vector defined by
u $[r]\ i.e\ dot\ product\ of\ u\ with\ row\ r\ of\ M$

3.3.2 Vector -Matrix Multiplication

In earlier section matrix-vector multiplication was discussed in terms of linear combinations of columns of a matrix. Next we see vector-matrix multiplication in terms of linear combinations of rows of a matrix.

Definition : In terms of Linear Combination :

Let M be RxC matrix over field F. Let w be a vector of dimension R . Then w ∗M is the linear combination $\sum_{r \in R} w\,[r](row\ r\ \ of\ M)$

Note : If M is R × C matrix but w is not of dimension R i.e it is not a R-vector then the product w ∗ M is illegal.

Example 3 : Suppose M = $\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \end{bmatrix}$ and w = $[1, 2]$

Then w ∗M can be computed since M is 2x3 and ν is 1x2 and result is :

w ∗ M = $\sum_{r \in R} w\,[r](row\ r\ \ of\ M)$

= 1[1, 0, 1] + 2[2 , 1, 3] = [1,0,1] + [4,2,6]

= [5,2,7]

Example 4 : Suppose M = $\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \end{bmatrix}$ and w = $[1, \ \ 0, \ 3]$

Then w ∗ M cannot be computed since M is 2x3 and ν is 1x3 and result is not valid (by note)

Next we will define vector- matrix multiplication in terms of dot product.

Definition : In terms of Dot Product:

Let M be RxC matrix over field F. Let u be a vector of dimension R. Then u*M is the C-vector defined by
u $[c]$ $i.e\ dot\ product\ of\ u\ with\ column\ c\ of\ M.$

Example 5 : Suppose M = $\begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}$ and w = $[2, -1]$

Then matrix- vector multiplication in terms of dot product is :

1st entry is dot product of row 1 [1,0] with w = [1,0].[2,-1] = 2-0 = 2

2nd entry is dot product of row 2 [2,1] with w = [2,1].[2,-1] = 4-1 = 3

3rd entry is dot product of row 3 [3,2] with w = [3,2].[2,-1] = 6-2 = 4

Hence finally M*w = [2, 3, 4]

Similarly, vector-matrix multiplication in terms of dot product can be carried out.

3.4 OTHER CONCEPTS

In the following sections we will see some concepts related to matrices.

3.4.1 Null Space

In earlier chapters we came across concept of homogeneous linear systems. It is the system where all values on right hand side of the equation are 0. We can define such a system as A*x = 0 i.e in the form of matrix-vector equation. In above equation right hand side of the equation is 0.

Definition : The null space of the matrix A is defined by the set

{v/ A*v = 0}. It is denoted by Null A

From the above definition it can be seen that null A is basically set of all solutions of homogeneous linear system, hence it also forms a vector space.

Example 6 : Suppose A = $\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$ then null(A) is all vectors such that

A*x = 0

i.e $\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} * \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ which gives $x_1 = 0$ and $x_1 + 2\,x_2 = 0$

hence Null(A) = {(0,0)}

Null(A) can also be computed easily using Row reduction form.

3.4.2 Computing Sparse Vector-Product

Definition : Sparse matrix is defined as a matrix whose most of the elements are 0.

In earlier sections we saw matrices in terms of vector and their products. For calculating product of matrices with vectors we can use either dot product or linear combinations definitions discussed earlier. But alone they cannot be conveniently used. Hence we combine both which leads to following definition :

Definition : Let M be RxC matrix over field F. Let u be a vector of dimension C. Then M * u is the vector v of dimension R, such that for each r ∈R, v[r] = $\sum_{c \in C} M\ [r, c] u[c]$

3.4.3 Linear Functions

Definition : Let U and V be vector spaces over a field F. Then a function

f: U → V is called a linear function if following properties are satisfied :

P1 : For any vector u ∈Domain(f) and α ∈ F is any scalar then

f(αu) = αf(u)

P2 : For any vectors u,v ∈Domain(f) then

f(u+v) = f(u) + f(v)

Linear function are called as linear transformation.

Let M be an R × C matrix over a field F, let $f : F^C \rightarrow F^R$ be defined by

by $f(x) = M * x$. Since the domain and co-domain are vector spaces,

function f satisfies Properties P1 and P2. Thus f is a linear function.

Example 7 : Let F be any field. Define function from F^2 to F by

$(x, y) \rightarrow x - y$ is a linear function.

P1 : For any vector $u = (x_1,y_1) \in F^2$ and $\alpha \in F$ be any scalar then consider

$f(\alpha u) = f(\alpha (x_1,y_1)) = f((\alpha x_1,\alpha y_1)) = \alpha x_1 - \alpha y_1 = \alpha (x_1 - y_1) = \alpha f(u)$

P2 : For any vectors $u =(x_1,y_1)$, $v= (x_2,y_2) \in F^2$ then

Consider $f(u+v) = f((x_1,y_1) + (x_2,y_2)) = f((x_1 + x_2, y_1+y_2))$

$= (x_1 + x_2) – (y_1+y_2) = (x_1 - y_1) + (x_2 – y_2) = f((x_1,y_1)) + f((x_2,y_2))$

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

Hence from P1 and P2 f is a linear function.

Result : Let U and V be vector spaces over a field F and $f: U \rightarrow V$ be a linear function, then f maps the zero vector of U to the zero vector of V

Such functions is called kernel.

Definition : Let U and V be vector spaces over a field F and $f: U \rightarrow V$ be a linear function then the set {v/f(v) = 0 } is called as kernel of f denoted by Ker f.

The result of linear function can be extended to n number of vectors.

3.4.4 Inner Product

Let u and v be two vectors of dimension D. Consider the "matrix-matrix product" u^Tv. The first matrix has one row and second matrix one column. By the dot-product definition of matrix-matrix multiplication, the product contains one single entry whose value is given by u.v

Example 8 : Suppose $A = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 14 \end{bmatrix}$

Since the final value of u^Tv is single entry it is called as inner product.

3.4.5 Outer Product

Next suppose u and v be two vectors not necessary of same domain.

Consider u^Tv : For each element of the domain u and each element of the domain of v, the s,t element of u^Tv is u[s]v[t].

Example 9 : Suppose $A = \begin{bmatrix} u \\ v \\ w \end{bmatrix} \begin{bmatrix} x & y \end{bmatrix} = \begin{bmatrix} ux & uy \\ vx & vy \\ wx & wy \end{bmatrix}$

This type of product is called the outer product of vectors u and v.

3.5 SUMMARY

This chapter gives different concepts of matrices and their examples. It will create base for the next concept of basis.

3.6 EXERCISE

1. Compute the following matrix-vector products

   a. M = $\begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -2 \end{bmatrix}$ and ν = $[2, \quad -3, \quad 0]$

   b. M = $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ and ν = $[2, \quad 4]$

2. For each of the following problems, answer whether the given matrix-matrix product is valid or not. If it is valid, give the number of rows and the number of columns of the resulting matrix (you need not provide the matrix itself).

   a. $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} 2 & -1 & 2 \\ 1 & 0 & -1 \end{bmatrix}$

   b. $\begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 & 1 \\ 1 & 1 & -1 \end{bmatrix}^T$

   c. $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}^T$

3. Compute Matrix Matrix Multiplication :

   a. $\begin{bmatrix} 2 & 2 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

   b. $\begin{bmatrix} 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ -2 & 6 \\ 1 & -1 \end{bmatrix}$

3.7 REFERENCES

- Coding the Matrix Linear Algebra through Applications to Computer Science Edition 1,PHILIP N. KLEIN, Newtonian Press (2013)

- Linear Algebra and Its Applications, Gilbert Strang, Cengage Learning, 4th Edition (2007).

❄❄❄❄❄❄❄

4

BASIS

Unit Structure:

4.0 OBJECTIVES

After going through this chapter, students will able to learn

- To understand spanning vectors
- To understand concept of Linear dependence and independence
- To learn concept of basis and dimension

4.1 INTRODUCTION

After learning the concepts of vector space, linear function in earlier chapters in this chapter we will learn concept of basis.

Basis has several properties which can be further used to justify concepts like linear dependence, independence, maximal linearly independent set etc.

The basis also tells us about the smallest set of vectors needed to span a vector space. Thus it helps to give information about structure of a vector space.

4.2 COORDINATE SYSTEM

A coordinate system is defined as a method for recognizing the location of a point. Most of the coordinate systems use two numbers i.e. a coordinate to detect a point or a location. These numbers indicate the distance between the point and some fixed point of reference called the origin.

For a vector space V in vector analysis, a coordinate system is indicated by a set of vectors $a_1, a_2, \ldots a_n$ of V such that every vector of the vector space can be written as linear combination of these vectors .

That is there exists scalars or real numbers $\alpha_1, \alpha_2, \ldots \alpha_n$ such that

$\nu = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_n a_n$ where $\nu \in V$ (any vector)

From discussion above the vector ν can be represented by $[\alpha_1, \alpha_2, \ldots, \alpha_n]$

of coefficients. These coefficients are called coordinates and the vector $[\alpha_1, \alpha_2, \ldots, \alpha_n]$ is called the coordinate representation of ν in terms of $a_1, a_2, \ldots a_n$. Also, this representation of ν is unique.

Example 1 : if we consider the vector [1, 3, 5, 2] it can be represented as :

[1, 3, 5, 2] = 1 [1, 0, 0, 0] + 3 [0, 1, 1, 0] + 2 [0, 0, 1, 1]

Hence the coordinate representation of ν in terms of [1, 0, 0, 0] , [0, 1, 1, 0] and [0, 0, 1, 1] is [1, 3, 5, 2]

4.3 TWO GREEDY ALGORITHMS FOR SET OF GENERATORS

Suppose we want to answer this question : For a given vector space V, what is the minimum number of vectors whose linear span is V?

To answer this, in this section we consider two algorithms

1. Grow algorithm

   def Grow(V)

   B = ϕ repeat while possible :

   Find a vector in V that is not in Span (B) and add it to B

   The algorithm halts when there is no more vector to add in B. By this time we can find the generating set.

   Example 2: Consider V = $\mathbb{R}^3$. In first iteration we add vector [1, 0, 0] to B . Next since [0, 0, 1] does not belong to Span(B) we add it to B. thus B = { [1, 0, 0], [0, 0, 1]}. Similarly in 3rd iteration we add [0, 1, 0] to B as it does not belong to span of B. Next if we consider any vector in $\mathbb{R}^3$

We can see it can be written as linear combination of either all or some of vectors of B. Hence there nr u o more vector to add to B, hence the algorithm stops.

2. Shrink algorithm

Exactly opposite to grow as name says we remove an element in every step.

def Shrink(V)

B = some finite set of vectors in V such that span(B) = V

repeat while possible :

Find a vector in V such that Span (B- {v}) = V and remove it from B

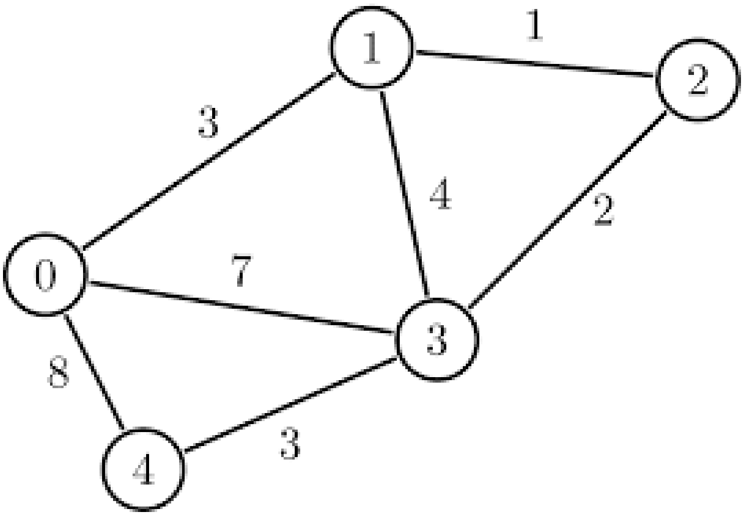
The algorithm halts when there is no more vector to remove from B such that spanning property is still satisfied. By this time we can find the generating set.

Example 2: Consider V = $\mathbb{R}^3$ and B ={ [1, 0, 0], [0, 0, 1],[0, 1, 0], [3, 2, 0], [0, 3, 1]}. In first iteration we remove vector [3, 2, 0] from B since [3, 2, 0] = 3[1, 0, 0] +2 [0, 1, 0]. Next we remove [0, 3, 1] as it belong to Span(B). Thus B = {[1, 0, 0], [0, 0, 1], [0, 1, 0]} . Now the algorithm stops since there is no more vector to remove.

4.4 MINIMUM SPANNING FOREST AND GF(2)

In this section we will see grow and shrink algorithm using graph theory that is minimum spanning problem.

Suppose we are given a graph with weights as below:



Suppose vertices represent cities and edges represent distances to travel from one city to another. Our goal is to travel from one city to another in covering all cities with minimum distance

To find minimum distance there are several algorithms but we will use grow and shrink algorithm

Grow algorithm

def Grow(G)

B = ϕ

Consider the edges in order from low to high

For each edge e:

If endpoint of e is not yet connected via edges add it to B

For above graph weights in increasing order are : 8 7 4 3 3 2 1

The solution obtained is 8 7 4 2

Shrink algorithm

def shrink(G)

B = { all edges }

Consider the edges in order from high to low

For each edge e:

If pair of nodes are connected via B – {e}:

Remove e from B

For above graph weights in increasing order are : 1 2 3 3 4 7 8

The solution obtained is 1 2 3 3 4

The Grow and Shrink algorithms for minimum spanning forest look like those algorithms used for finding a set of generators for a vector space.

In this section, we describe how to model a graph by means of vectors over GF(2).

Let C = {set of vertices of graph} = {0,1,2,3,4} be the set of nodes

A subset of C is characterized by the vector with ones in the corresponding entries and zeroes elsewhere.

A subset of C is represented by the vector with ones in the corresponding entries and zeroes elsewhere.

Hence the vectors corresponding to all the edges in our graph are :

Edge	Vector				
	0	1	2	3	4
{0,4}	1				1
{0,3}	1			1	
{1,3}		1		1	
{3,4}				1	1
{1,2}		1	1		
{2,3}			1	1	

In general, a vector with 1's in entries x and y is the sum of vectors corresponding to edges that form an x-to-y path in the graph. Thus, for these vectors, it is easy to tell whether one vector is in the span of some others.

4.5 LINEAR INDEPENDENCE

Lemma (Superfluous-Vector Lemma): For any set S and any vector $v \in S$, if v can be written as a linear combination of the other vectors in S then Span $(S-\{v\})$ = Span S

Definition: Let V be a vector space .Then vectors $v_1, \ldots, v_n$ in V are called as linearly dependent if the zero vector can be written as a nontrivial linear combination of these vectors. That is

$0 = \alpha_1 v_1 + \cdots + \alpha_n v_n$

Here we denote the linear combination as a linear dependency in $v_1, \ldots, v_n$.

Example .3: The vectors [1, 0, 0], [0, 3, 0], and [3, 9, 0] are linearly dependent, as shown by the following equation:

3 [1, 0, 0] + 3 [0, 3, 0] − 1 [3, 9, 0] = [0, 0, 0]

Thus 3 [1, 0, 0] + 3 [0, 3, 0] − 1 [3, 9, 0] is a linear dependency in [1, 0, 0], [0, 3, 0], and [3, 9, 0].

Example 4: The vectors [1, 0, 0], [0, 3, 0], and [0, 0, 5] are linearly independent.

Since if we consider α_1 [1, 0, 0] + α_2 [0, 3, 0] + α_3 [0, 0, 5] = [0, 0, 0]

Then all scalars $\alpha_1, \alpha_2, \alpha_3$ all are 0.

Properties of linear (in)dependence

1. A subset of a linearly independent set is linearly independent.
2. Let $v_1, \dots, v_n$ be vectors. A vector v_i belongs to the span of the other vectors if and only if the zero vector can be written as a linear combination of $v_1, \dots, v_n$ in which the coefficient of v_i is nonzero.
3. The vectors obtained by the Grow algorithm are linearly independent.
4. The vectors obtained by the Shrink algorithm are linearly independent.

4.6 BASIS

In earlier sections we saw the Grow algorithm and the Shrink algorithm where each of them finds a set of vectors spanning the vector space V. In addition in each case, the set of vectors found is linearly independent.

Next we define basis of vector space one of the most important concept in linear algebra.

Definition: Let V be a vector space. A basis for V is a linearly independent set of generators for V.

In other words, a set B of vectors of V is a basis for V if B satisfies two properties:

PB1 Span B = V, (Spanning) and

PB2 B is linearly independent. (Independent)

Example 5: Let V the vector space spanned by [1, 0, 0], [0, 1, 1], and

[1, 1, 1].

Then the set {[1, 0, 0], [0, 1, 1], [1, 1, 1]} is not a basis for V because it is

not linearly independent as [1, 1, 1] = [1, 0, 0] + [0, 1, 1]

However, the set {[1, 0, 0], [0, 1, 1]} is a basis as it satisfies the above two properties.

Lemma : The standard generators for F^D form a basis.

Lemma (Unique-Representation Lemma): Let V be a vector space and B be a basis of V, then every vector in V can be uniquely represented as linear combination of vectors of B.

i.e Let $B = \{a_1, \dots, a_n\}$ be a basis for a vector space V. For any vector $v \in V$, there is exactly one representation of v in terms of the basis vectors.

4.7 DIMENSION

After defining basis in earlier section lets now see the number of elements in any given basis. Before that let us see some results with respect to basis.

Lemma (Morphing Lemma): Let V be a vector space. Suppose S is a set of generators for V, and B is a linearly independent set of vectors belonging to V. Then $|S| \geq |B|$.

Theorem (Basis Theorem): Let V be a vector space. All bases for V have the same size.

Theorem : Let V be a vector space. Then a set of generators for V is a smallest set of generators for V if and only if the set is a basis for V.

Definition : Let V be a vector space. Then the dimension of V is defined to be the size of a basis for V.

The dimension of a vector space V is written *dim V.*

If we consider example 5 then dim V = 2 since it has basis B containing 2 vectors i.e.[1, 0, 0] and [0, 1, 1]

Example 6: One basis for $\mathbb{R}^3$ is the standard basis:

{[1, 0, 0], [0, 1, 0], [0, 0, 1]}. Hence the dimension of R3 is 3.

- **4.7.1 DIMENSION AND RANK**

Definition : Rank of a set S of vectors is defined as the dimension of Span S.

We denote *rank S* for the rank of S.

Proposition : For any set S of vectors, rank $S \leq |S|$.

Definition : For a matrix M, the row rank of M is defined as the rank of its rows, and the column rank of M is defined as the rank of its columns.

Definition : For a matrix M, the row rank of M is the dimension of Row M, and the column rank of M is the dimension of Col M.

Example 7 : Consider the matrix

$$M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

Here row vectors are {[1, 0], [0, 1], [1, 1]} which are linearly dependent .but if we remove [1, 1] then vectors become independent. Hence Row rank = 2

Similarly column vectors are {[1, 0, 1] ,[0,1,1]} which are linearly independent as discussed earlier. Hence column rank = 2

In any case we have Row Rank = Column Rank

Definition : The rank of a matrix is defined to be its common value of column rank which is equal to its row rank.

Lemma (Superset-Basis Lemma): For any vector space V and any linearly independent set B of vectors, V has a basis that contains all of B.

The Dimension Principle

Using the Superset-Basis Lemma we can prove the following principle.

Lemma (Dimension Principle): If V is a subspace of vector space W then

PD1: dim V $\leq$ dim W, and

PD2: if dim V = dim W then V = W.

Example 8: Suppose W = Span {[1, 0], [1, 1]}. Clearly V is a subspace of $\mathbb{R}^2$. However, the set {[1, 0], [1, 1]} is linearly independent, so dim V = 2. Since dim $\mathbb{R}^2 = 2$, hence by PD2 V = $\mathbb{R}^2$.

- **4.7.2 DIRECT SUM**

We are acquainted with the idea of adding vectors—now we study about adding of vector spaces. These ideas will be advantageous in proving a fundamental theorem in the next section—the Kernel-Image Theorem.

Let U and V be two vector spaces consisting of D-vectors over a field F.

Definition : If U and V have only the zero vector in common then we define the direct sum of U and V to be the set $\{u + v : u \in U, v \in V\}$

We write direct sum of U and V as $U \oplus V$

That is, $U \oplus V$ is the set of all sums of a vector in U and a vector in V.

Example 9 : Let U = span{[1,0]} i.e X-axis and V = span{[0,1]} i.e Y-axis

Then $U \oplus V = \mathbb{R}^2$

Result : The direct sum $U \oplus V$ is a vector space.

Lemma : The set of generators for $V \oplus W$ is the union of a set of generators of V and a set of generators of W

Lemma (Direct Sum Basis Lemma): The union of a basis of U and a basis of V is a basis of $U \oplus V$.

Corollry :Any vector in $U \oplus V$ has a unique representation as u + v where $u \in U, v \in V$.

Definition : U and V are said to be complementary subspaces of W, if

$U \oplus V = W$

- **4.7.3 DIMENSION AND LINEAR FUNCTION**

In this section we will see how dimension can be related to linear functions studied in earlier sections. We will devise a criterion for invertibility of a linear function. That in turn will provide a criterion for matrix invertibility. These criteria will construct an important theorem, the Kernel-Image Theorem.

We have studied earlier that linear function $f : V \rightarrow W$ is invertible if

(i) f is one-to-one and (ii) f is onto.

By the One-to-One Lemma, we know that f is one-to-one iff its kernel is trivial.

Similarly there is a criterion for checking if a linear function is onto.

Recall : image of f is $\text{Im } f = \{f(v) : v \in V\}$. Thus f is onto iff $\text{Im } f = W$.

Also Im f is a subspace of W.

By the Dimension Principle, f is onto iff $\dim \text{Im } f = \dim W$.

Hence We can conclude:

A linear function $f : U \rightarrow W$ is invertible if $\dim \text{Ker } f = 0$ and $\dim \text{Im } f = \dim W$.

<u>The Kernel-Image Theorem</u>

For any linear function $f : V \rightarrow W$, $\dim \text{Ker } f + \dim \text{Im } f = \dim V$

Theorem (Linear-Function Invertibility Theorem): Let $f : V \rightarrow W$ be a linear function. Then f is invertible if and only if $\dim \text{Ker } f = 0$ and $\dim V = \dim W$.

Theorem (Rank-Nullity Theorem): For any n-column matrix A,

rank A + nullity A = n

Example 10

Let $T: \mathbb{P}_1 \rightarrow \mathbb{R}$ be the linear transformation defined by $T(p(x))=p(1)$ for all $p(x) \in \mathbb{P}_1$. Find the kernel and image of T, Verify the kernel–Image theorem.

We will first find the kernel of T : It consists of all polynomials in $\mathbb{P}_1$ that have 1 for a root.

$\ker(T)=\{p(x)\in\mathbb{P}_1 \mid p(1)=0\}=\{ax+b \mid a,b\in R \text{ and } a+b=0\}=\{ax-a \mid a\in R\}$

Therefore a basis for ker(T) is $\{x-1\}$ and dimension = 1

Notice that this is a subspace of $\mathbb{P}_1$.

Now consider the image. It consists of all numbers which can be obtained by evaluating all polynomials in $\mathbb{P}_1$ at 1.

im(T)={p(1) | p(x)∈P1}={a+b | ax+b∈P1}={a+b | a,b∈R} = $\mathbb{R}$

Therefore a basis for im(T) is {1} and dimension is 1

Dim($\mathbb{P}_1$) = 2 = 1+1 = dim(ker T) + Dim (im T)

Hence Kernel-Image theorem verified.

4.8 THE ANNIHILATOR

Definition : For a subspace V of F^n, the annihilator of V, denoted as V^o, is defined as $V^o = \{u \in F^n : u \cdot v = 0$ for every vector $v \in V\}$

Results :

1. Let $a_1, \ldots, a_m$ be generators for V, and let A = $[a_1, a_2, \ldots, a_m]^T$ Then V^o = Null A.

2. (Annihilator Dimension Theorem): Let V and V^o be subspaces of F^n, where F is a field , then dim V + dim V^o = n

3. (Annihilator Theorem): $(V^o)^o = V$ (The annihilator of the annihilator is the original space.)

4.9 SUMMARY

In this chapter we studied about basis of a vector space, its dimension and their properties .

4.10 EXERCISE

1. Let V = Span {[0, 0, 1], [1, 0, 1], [2, 1, 1]}. For each of the following vectors, show it belongs to V by writing it as a linear combination of the generators of V.

   (a) [2, 1, 4]

   (b) [1, 1, 1]

   (c) [5, 4, 3]

   (d) [0, 1, 1]

2 Let V = Span {[0, 1, 0, 1], [0, 0, 1, 0], [1, 0, 0, 1], [1, 1, 1, 1]} where the vectors are over GF(2). For each of the following vectors over GF(2), show it belongs to V by writing it as a linear combination of the generators of V.

(a) [1, 1, 0, 0]

(b) [1, 0, 1, 0]

(c) [1, 0, 0, 0]

3 For each of the set given below, show the given vectors over R are linearly dependent.

(a) [1, 2, 0], [2, 4, 1], [0, 0, −1]

(b) [2, 4, 0], [8, 16, 4], [0, 0, 7]

(c) [0, 0, 5], [1, 34, 2], [123, 456, 789], [−3, −6, 0], [1, 2, 0.5]

4 For each of the following matrices, (a) give a basis for the row space (b) give a basis for the column space, and (c) verify that the row rank equals the column rank. Justify your answers.

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

5 Verify Rank – Nullity theorem

(a) $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by T(x, y) = x+y

(b) $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by T(x, y) = (x, x+y, y)

4.11 REFERENCES

- Coding the Matrix Linear Algebra through Applications to Computer Science Edition 1,PHILIP N. KLEIN, Newtonian Press (2013)
- Linear Algebra and Its Applications, Gilbert Strang, Cengage Learning, 4th Edition (2007).

❄❄❄❄❄❄❄

5

GAUSSIAN ELIMINATION

Unit Structure:

5.0 OBJECTIVES

After going to this chapter, you will be able to:

i.) Solve a set of simultaneous linear equations using Gauss elimination,

ii.) Perform elementary row operations to produce zeros below the diagonal of the coefficient matrix to reduce it to echelon form.

iii.) Find basis for the null space.

5.1 INTRODUCTION

Given a linear system expressed in matrix form AX = B, where A is coefficient matrix and X is variable matrix. Gaussian elimination method is used to solve a system of linear equations by performing elementary row operations. Elementary row operations are categorized as: a.) Interchange any two rows; b.) Multiply a row by a nonzero constant; c.) Add a multiple of one row to another row. This row reduction algorithm continues till we get 0s (i.e., zeros) on the lower left-hand corner of the matrix as much as possible. That means the obtained matrix should be an upper triangular matrix.

5.2 ECHELON FORM

Pivot: A pivot is the first non-zero element in a row and leading coefficient in a column with all the rows below containing 0's.

Echelon Form of a matrix: There are two types of Echelon form of a matrix:

i.) Row Echelon form: A matrix is said to be in row echelon form (ref) when it satisfies the following conditions:

- The first non-zero element is 1.

- Each leading entry is in a column to the right of the leading entry in the previous row.
- Rows with all zero elements, if any, are below rows having a non-zero element.

ii.) Reduced row Echelon form: A matrix is said to be in reduced row echelon form (ref) when it satisfies the following conditions:

- The matrix is in its row echelon form.
- The leading entry in each row is the only non-zero entry in its column.

Uses of Echelon form:

- If a matrix is in echelon form, the non-zero rows form a basis for the row space

  Example: $A = \begin{bmatrix} 2 & 3 & 1 & 0 \\ 0 & 4 & 0 & 1 \\ 0 & 0 & 9 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ then the rows [2 3 1 0], [0 4 0 1] and [0 0 9 6] are the basis of the row space.

- If an echelon form of a matrix has neither pivots in all rows nor all columns, the given set of vectors are linearly dependent.

  let V = {(1, 1, 1), (1, 2, 3), (1, 4, 7)}

  we compute $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$

  since A has neither pivots in all rows nor in all columns, the set is linearly dependent.

- The number of non-zero rows in row echelon form of a matrix is equal to rank of the matrix.

  Example $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 5 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$ [by performing elementary row operations]

  The number of non-zero rows = 2, hence the rank of the matrix A = 2.

5.3 GAUSSIAN ELIMINATION OVER GF(2)

Gaussian elimination is very simple process for matrices over GF(2). The required row operations consist only XOR of two rows and swapping of two rows. Solving linear systems over GF(2) is of particular interest in cryptography and crypto-analysis.

The Gaussian elimination over GF(2) on a matrix A requires elementary column operations rather than elementary row operations.

Let us take an example:

Let Q = {6, 42, 105, 20, 63} and P = {2, 3, 5, 7}

We have,

$6=2^1 3^1 5^0 7^0$

$42=2^1 3^1 5^0 7^1$

$105=2^0 3^1 5^1 7^1$

$20=2^2 3^0 5^1 7^0$

$63=2^0 3^2 5^0 7^1$

We define A as $A=\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{bmatrix}$ (mod 2)

⇨ $A=\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Performing elementary column operations and mark each row which has a point

Since $A_{12} = 1$, and $c_2 \longrightarrow c_2 + c_1$, we get

⇨ $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Again performing $c_3 \longrightarrow c_3 + c_2$, and $c_4 \longrightarrow c_4 + c_2$, we get

⇨ $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Now performing $c_1 \longrightarrow c_1 + c_4$, we get

⇨ $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$

Note that row 5 has not been used, since $A_{51} = A_{54} = 1$, row 5 and all rows for which $A_{i1} = 1$ and $A_{i4} = 1$ are dependent. From the above matrix we see that rows 1, 2, and 5 are dependent. If we sum row1, row2, and row5 in GF(2), we obtain a zero row.

i.e. 1 0 0 0 Row $1(Q_1 = 6)$

0 0 0 1 Row $2(Q_2 = 42)$

1 0 0 1 Row $5(Q_5 = 63)$

———

0 0 0 0

This implies that R={Q_1, Q_2, Q_5} and product $Q_1Q_2Q_5$ forms perfect square.

$Q_1Q_2Q_5$ = 6*42*63 = 126^2

5.4 SOLVING A MATRIX-VECTOR EQUATION USING GAUSSIAN ELIMINATION

Consider a system of linear equation of n unknowns and n equations as

$$a_{11}x_1 + a_{12}x_2 + \ldots\ldots\ldots\ldots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \ldots\ldots\ldots\ldots + a_{2n}x_n = b_2$$

.

.

.

$$a_{n1}x_1 + a_{n2}x_2 + \ldots\ldots\ldots\ldots + a_{nn}x_n = b_n$$

Step 1: To eliminate x_1 from second, third,……n^{th} equations:

Assuming $\mathbf{a_{11}} \neq 0$, we eliminate x_1 from the second equation by subtracting $\mathbf{a_{21}/a_{11}}$ times the first equation from the second equation.

Similarly we eliminate $\mathbf{x_1}$ from the third equation by subtracting $\mathbf{a_{31}/a_{11}}$ times the first equation from the third equation.

By proceeding in the similar way, we get the following new system of equations as,

$$a_{11}x_1 + a_{12}x_2 + \ldots\ldots\ldots\ldots + a_{1n}x_n = b_1$$

$$a_{22}'\, x_2 + \ldots\ldots\ldots\ldots + a_{2n}'x_n = b_2'$$

.

.

.

$$a_{n2}'\, x_2 + \ldots\ldots\ldots\ldots + a_{nn}'x_n = b_n'$$

From the above it is clear that, the first equation is called pivotal equation and a_1 is called first pivot.

Step 2: To eliminate x_2 from the third equation:

Assuming $a_{12}' \neq 0$, we eliminate x_2 from third equation by subtracting (a_{32}'/a_{22}') times the second equation from the third equation. Thus we get the following new system as,

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \dots\dots\dots\dots + a_{1n}x_n &= b_1 \\
a_{22}'x_2 + \dots\dots\dots\dots + a_{2n}'x_n &= b_2' \\
&\;\;\vdots \\
+ \dots\dots\dots\dots\dots\dots + a_{nn}''x_n &= b_n''
\end{aligned}
$$

Step 3: To evaluate the unknowns:

The values of unknowns $x_1, x_2, \dots\dots\dots\dots x_n$ are found from the above reduced system by back substitution.

Gauss Elimination Method

Example 1: Solve the following system of equations by Gaussian elimination method:

2x + y + z =10; 3x + 2y + 3z =18; x + 4y + 9z = 16

Solution:

2x + y + z =10 ------------------(i)

3x + 2y + 3z =18----------------(ii)

x + 4y + 9z = 16------------------(iii)

Multiplying equation (iii) by 2

2x + 8y + 18z = 32-------(v)

Subtracting equation (i) from (iv)

7y + 17z = 22

Performing 7 * (ii) + (v) we get,

2x + y + z = 10----------(i)

y + 3z = 6-----------(iv)

4z = 20 ------------(vi)

from equation(vi), we get $z = \frac{20}{4} = 5$

using back substitution method, we get, y = -9 and x = 7.

$\therefore$ x = 7, y = -9 and z = 5

Example 2: Solve the following system of equations by Gaussian elimination method:

y-z = 3: -2x + 4y – z =1: and -2x + 5y – 4z = -2

Solution: Consider

-2x + 4y – z =1----------------(i)

-2x + 5y – 4z = -2--------------(ii)

y-z = 3----------------------------(iii)

Subtracting equation (ii) from equation (i), we get

-2x + 4y – z =1-----------------(i)

-y + 3z = 3-----------------(iv)

y – z = 3-----------------(iii)

Adding equation (iii) and equation (iv), y-z+-y + 3z = 3+3

2z = 6 ⇒z = $\frac{6}{2} = 3$ ⇒ z = 3

Substituting z =3 in equation (iv),

-y + 3(3) = 3

⇒ -y = 3 – 9 ⇒ -y = - 6 ⇒ y =6

Substituting y = 6 and z = 3 in equation (i),

⇒ -2x + 4(6) – 3 = 1

⇒ -2x = 1 – 24 + 3 ⇒ -2x = -20 ⇒ x = $\frac{-20}{-2} = 10$

The solution of the given set of equations are x = 10, y = 6 and z = 3.

Example 3: Solve the following system of equations by Gaussian Elimination method:

5x + 4y - z = 0; 10y – 3z = 11; z = 3;

Solution: Given the system of equations are,

5x + 4y – z = 0------------------(i)

10y - 3z = 11--------------(ii)

z =3----------------(iii)

Performing back substitution, z = 3.

Putting value of z in equation (ii), We get,

10y - 3(3) = 11 ⇒ 10y = 11 + 9 ⇒ 10y = 20 ⇒ y = $\frac{20}{10} = 2$

Substituting values of y and z in equation (i),

5x + 4(2) – 3 = 0 ⇒ 5x = 3 – 8 ⇒ 5x = -5 ⇒ x = $\frac{-5}{5}$ = -1

∴ x = -1, y = 2 and z = 3.

5.5 FINDING A BASIS FOR THE NULL SPACE

This topic explains you how to find the basis for the null space of a mxn matrix A using Gaussian Elimination method.

We have A•X = 0, either the solution is unique and X = 0 is the only solution or there are infinitely many solutions, which can be parametrized by non-pivotal elements.

The basis of a null space of a matrix A is defined as Null (A) ={V: A•V = O}. The dimension of the null space of A is called nullity of A.

To find basis for the null space, we convert the coefficient matrix into row echelon form.

Example 1: Let A= $\begin{bmatrix} -4 & -1 & -3 & -2 \\ 0 & 4 & 0 & -1 \end{bmatrix}$. Find basis for the null space of A.

Solution: Let X= $\{(x_1, x_2, x_3, x_4): A\bullet X = O\}$ is a basis for the null space of A.

Then $\begin{bmatrix} -4 & -1 & -3 & -2 \\ 0 & 4 & 0 & -1 \end{bmatrix} . \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = O$

Matrix A is in row echelon form:

Hence $-4x_1 - x_2 - 3x_3 - 2x_4 = 0$ ——————————————(i)

and $4x_2 - x_4 = 0$ ——————————————————(ii)

$\Rightarrow x_4 = 4x_2$

Substituting x_4 equation (i), we get

$-4x_1 - 9x_2 - 3x_3 = 0 \Rightarrow 3x_3 = -4x_1 - 9x_2$

Writing vector components x_1, x_2, x_3 and x_4 in the following manner,

$$\begin{array}{r} x_1 = 1x_1 + 0x_2 \\ x_2 = 0x_1 + 1x_2 \\ x_3 = \frac{-4}{3}x_1 + \frac{-9}{3}x_2 \\ x_4 = 0x_1 + 4x_2 \end{array} = x_1 \begin{bmatrix} 1 \\ 0 \\ \frac{-4}{3} \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ \frac{-9}{3} \\ 4 \end{bmatrix}$$

Since x_1 and x_2 are arbitrary, the basis of null space of A is span of {(1, 0, $\frac{-4}{3}$, 0), (0, 1, $\frac{-9}{3}$, 4)}.

5.6 FACTORING INTEGERS

The unique factorization theorem: Every positive integer a>1 can be expressed uniquely as a product of positive primes.

To find a nontrivial factor of a composite number n is the main concern. The simplest factoring algorithm is the trial division method which tries all the possible divisors of n to complete prime factorization:

n = $p_1 p_2 \ldots\ldots p_r$

Algorithm for factoring integer n by trial divisions:

[1] Input n and set r $\leftarrow$ 0, k $\leftarrow$ 2.

[2] If n = 1, go to step [5].

[3] q$\leftarrow$ n/k and t $\leftarrow$ n (m0d k).

If t ≠ 0. Go to [4].

r$\leftarrow$r+1, $p_r \leftarrow$k, n$\leftarrow$q, go to [2].

[4] If q > k, then k$\leftarrow$k+1, and go to [3].

r$\leftarrow$r+1, $p_r \leftarrow$n.

[5] Exit; terminate the algorithm.

An improvement of algorithm is to make use of an auxiliary sequence of trial divisors:

$2 = d_0 < d_1 < d_2 < d_3 < d_4 < \ldots$ which includes all primes $\sqrt{n}$ and at least one value $d_k \geq \sqrt{n}$.

<u>The number of divisors of a positive integer:</u> Let n is a positive integer such that n>1. Then by unique factorization theorem, n can be expressed as product of positive primes.

Let n = $p_1^{\alpha_1} p_2^{\alpha_2} \ldots\ldots\ldots p_r^{\alpha_r}$where 1<$p_1 < p_2 < \ldots\ldots < p_r$ and p's are positive primes and $\alpha_1 \alpha_2 \ldots\ldots\ldots \alpha_r$ are positive integers. Then the number of distinct positive integral divisors of n=$(1+\alpha_1)(1+\alpha_2)\ldots\ldots\ldots\ldots\ldots(1+\alpha_r)$ and it is denoted by T(n).

Also the sum of all the terms in the product:

P =

$$\left(\frac{p_1^{\alpha_1+1}-1}{p_1-1}\right)\bullet\left(\frac{p_2^{\alpha_2+1}-1}{p_2-1}\right)\bullet\ldots\ldots\ldots\ldots\bullet\left(\frac{p_r^{\alpha_r+1}-1}{p_r-1}\right)$$ and it is denoted by $\sigma(n)$.

<u>Greatest Common Divisor:</u> For a,b ∈Z, the largest d ∈Z, which divides both a and b, is called greatest common divisor of a and b .

Let d = gcd(a, b)

Each common divisor d of a and b divides gcd(a, b) .

If gcd(a, b) =1, we call a and b coprime.

The gcd of a and b has a representation.

gcd(a, b) = x•a+y•b, with integers x, y ∈ Z.

If gcd(a, b) = 1. Then $\bar{a}$ is called primitive residue class modulo n.

Euclidian Algorithm: Euclidian algorithm enables us to find the actual value of the greatest common divisor d of two given integers a and b and also to find integers x and y such that

d = x•a+y•b

Example: Find (26,118) and express it in the form 26x+118y, where x and y ∈Z.

Solution: We have,

118 = 26*4+14

⇒26 = 14*1+12

⇒14 = 12*1+2

⇒12 = 2*6+0

Hence the last non-zero remainder is 2 = (26, 118).

From the last we get,

2 = 14-12*1 = 14-12

⇒ 2 = 14 - (26 - 14) = 2*14 – 26

⇒ 14 = 118 - (26)*4

⇒2 = 2[118 - (26)*4] – 26

⇒2 *118 – 9*26-------------------(i)

Hence (26, 118) = 2

Equation(i) is in the form of 26x+118y, by comparison, we get,

x = 9 and y = 2

Example 2: Find the number of distinct positive integral divisors and their sum for the integers 56700.

Solution: Expressing 56700 as a product of prime integers as,

$56700=2^2 * 3^4 * 5^2 * 7$

Here $p_1 = 2, p_2 = 3, p_3 = 5, p_4 = 7, \alpha_1 = 2, \alpha_2 = 4, \alpha_3 = 2, \alpha_4 = 1$

Then, number of distinct positive integral divisors of 56700 is

T(56700) = (2+1) (4+1) (2+1) (1+1) = 90

And the sum of all distinct positive integral divisors

$\sigma(56700) = \frac{2^{2+1}-1}{2-1} * \frac{3^{4-1}-1}{3-1} * \frac{5^{2+1}-1}{5-1} * \frac{7^{1+1}-1}{7-1}$ =7*121*31*8 = 210056.

5.7 SUMMARY

Any matrix can be transformed to reduced row echelon form by using Gaussian elimination method. This is particularly useful for solving systems of linear equations. The echelon form of a matrix isn't unique, which means there are infinite answers possible after performing row reduction. But the reduced row echelon form is unique, which means row-reduction on a matrix will produce the same answer no matter how you perform the same row operations. The method can be applied even if the coefficient matrix is singular matrix or rectangular matrix. Gaussian elimination is also needed to determine the rank of a matrix.

5.8 REFERENCES

Linear Algebra and its Applications, David C Lay, Pearson Education India; 3rd Edition, 2002.

<u>Exercise</u>

Q.1: Solve the following system of linear equations by Gaussian-Elimination method:

i.) $x + y = 3$ and $3x - 2y = 4$

ii.) $x + y + z = 3$; $2x + 3y + 4z = 9$; $x - 2y + 3z = 2$

iii.) $x + y - z = 9$; $-x - 2z = 2$; $y + 3z = 3$

Q. 2: Find the basis for null spaces of the following matrices:

i.) $\begin{bmatrix} 1 & 0 & 3 & 2 & 1 \\ 0 & 2 & 2 & 4 & 4 \\ 0 & 0 & 0 & 6 & 6 \end{bmatrix}$

ii.) $\begin{bmatrix} 0 & 0 & 0 & -1 \\ 4 & -1 & 1 & -1 \\ 8 & -2 & 3 & -1 \end{bmatrix}$

❄❄❄❄❄❄❄

6

INNER PRODUCT AND ORTHOGONALITY

Unit Structure:

6.0 OBJECTIVES:

After going to this chapter, you will be able to:

- Find inner product of two vectors.
- Determine whether the given vectors are orthogonal to each other or not.
- Construct orthogonal set of generators.
- Find orthogonal complement of any vector v.

6.1 INNER PRODUCT:

Let u = (u_1 , u_2, … … … … … . , u_n) and v = (v_1 , v_2, … … … … … . , v_n) are two n-vectors of a real vector space. The inner product of u and v is given by the sum of the products of the coordinates with same index. It is also defined as the dot product of corresponding components of u and v. It is denoted as <u, v>.

$<u, v> = u_1v_1 + u_2v_2 + \ldots\ldots\ldots + u_nv_n$.

The inner product of two vectors satisfies the following properties:

i. $< u, u > \geq 0 \Rightarrow < u,u > = 0$ iff u = 0

ii. $< u, v > = < v, u >$ (symmetry)

iii. $< u + w, v > = < u, v > + <w, v>$ (linearity)

iv. $< u, w + v > = <u, w > + <u, v>$ (linearity)

v. $< \alpha u, v > = \alpha < u, v>$ (homogeneity)

Any linear space that satisfies the above postulates is called inner product space.

6.1.1 Norm of a Vector:

The norm of a vector v ∈ V is defined as the positive square root of the inner product of the vector with itself. The norm of a vector v is written as ||v||.

$\| v \| = \sqrt{< v, v >} = \sqrt{{v_1}^2 + {v_2}^2 + \cdots + {v_n}^2}$

6.1.2 Norm of distance of two vectors:

Norm of distance between two vectors u and v is defined as d(u, v) = || u-v || = < u-v, u-v >

$= \sqrt{(u-v)\bullet(u-v)} \quad = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}$

Example 1: If u = (1, -3, 5) and v = (3, 1, -4), find the inner product of u and v. Also find norm of u, norm of v, and norm of distance between u and v.

Solution: Inner product of u and v = <u, v>

= 1*3 + (-3)*1 + 5*(-4) = 3 – 3 – 20 = -20

Norm of u = $\sqrt{1^2 + (-3)^2 + 5^2} = \sqrt{35}$

Norm of v = $\sqrt{3^2 + 1^2 + (-4)^2} = \sqrt{26}$

Norm of distance between u and v =
$\sqrt{(1-3)^2 + (-3-1)^2 + (5-(-4))^2} = \sqrt{4+16+81} = \sqrt{101}$

<u>Theorem 1: Cauchy-Schwartz inequality:</u>

For any vectors u,v in an inner product space v, $< u, v >^2 \leq$<u,v><u,v> or |<u,v>| ≤ ||u|| ||v||.

Proof: Let y = y(t) = <u + tv, u + tv>, t∈R

= <u, u+tv> + <tv, u + tv> (by linearity)

= <u, u> + 2<u, v> t + <v, v> t^2

It is a quadratic equation.

⇒ <u, u> + 2<u, v> t + <v, v> t^2 = 0

It has at most one solution as y(t) ≥ 0. This implies that its discriminant must be less or equal to zero.

i.e. $[2 < u, v >]^2 - 4 < u, u >< v, v > \leq 0$

⇒ $4(< u, v >)^2 \leq$ 4<u, u> <v, v>

⇒ $(< u, v >)^2 \leq$ <u, u> <v, v>

or $|< u, v >| \leq \|u\| \, \|v\|$

Hence proved.

Note: For non-zero vector u,v ∈ V , the Cauchy-Schwartz inequality implies that

$$-1 \leq \frac{<u,v>}{||u||||v||} < 1$$

The angle θ between u and v is defined by cos θ $= \frac{<u,v>}{||u||||v||}$ **the angle is unique.**

6.2 ORTHOGONALITY:

The two vectors u and v are orthogonal, if they are perpendicular to each other. In other words, the two vectors are said to be orthogonal to each other if angle between them is 90°.

In terms of inner product, we can define that two vectors are orthogonal if their inner product is equal to zero.

<u>Orthogonal sets:</u> A set S $= \{u_1, u_2, \dots, u_n\}$ of non-zero vectors of V is called an orthogonal set if every pair of vectors are orthogonal to each other.

i.e. $< u_i, u_j > = 0$, $1 \leq i < j \leq n$.

This orthogonal set of vectors becomes orthonormal if in addition $< u_i, u_i >$ = 1 for all i ≤ n.

Theorem 2 : Pythagorean Theorem: Let $v_1, v_2, \dots, v_n$ be mutually orthogonal vectors. Then,

$$||v_1 + v_2 +, \dots + v_n||^2 = ||v_1||^2 + ||v_2||^2 + \cdots + ||v_n||^2$$

Proof: Let n=2,

If u and v are orthogonal, then <u, v> = 0

$\Rightarrow ||u+v||^2 =$ <u + v, u + v> = <u, u> + <u, v> + <v, u> +<v, v>

= <u, u> + 2 <u, v> + <v, v> (by symmetry)

=<u, u> + <v, v> (u and v are orthogonal)

$= ||u||^2 + ||v||^2$

Similarly we can prove that

$$||v_1 + v_2 +, \dots + v_n||^2 = ||v_1||^2 + ||v_2||^2 + \cdots + ||v_n||^2.$$

Example 1: Determine if u = (3, 2, 0, -5) and v = (-4, 1, 6, -2) are orthogonal.

Solution: If <u, v> = 0, the two vectors u and v are orthogonal.

<(3, 2, 0, -5), (-4, 1, 6, -2)> = 3*(-4) + 2*1 + 0*6 + (-5)*(-2) = 0.

Hence, vectors u and v are orthogonal.

Example 2:

Verify Pythagorean theorem for u = (1, 0, 2, -4) and v = (0, 3, 4, 2)

Solution: Pythagorean theorem for u and v is $||u + v||^2 = ||u||^2 + ||v||^2$

Consider, L.H.S:$||u + v||^2$ = < u + v, u + v>

we have u+v = (1, 0, 2, -4) + (0, 3, 4, 2) = (1, 3, 6, -2)

$||u + v||^2$ = <(1, 3, 6, -2), (1, 3, 6, -2)> = 1 + 9 + 36 + 4 = 50

consider R.H.S: $||u||^2 + ||v||^2$ = <u, u> +<v, v>

= <(1,0,2,-4),(1,0,2,-4)> + <(0,3,4,2),(0,3,4,2)>

=21+29

=50

∴L.H.S = R.H.S

Hence Proved.

Example 3: Find inner product, angle, orthogonality for

p = -5+2x-x^2 and q = 2+3x^2.

Solution: Let u = (-5, 2, -1) and v = (2, 0, 3)

Inner product of p and q is <u, v> = -5*2 + 2*0 + (-1)*3 = -10 + 0 - 3= -13

||u|| = $\sqrt{(-5)^2 + 2^2 + (-1)^2} = \sqrt{30}$

||v|| = $\sqrt{2^2 + 0 + 3^2} = \sqrt{13}$

Angle between p and q is cos θ $= \frac{<u,v>}{||u||\,||v||} = \frac{-13}{\sqrt{30}\sqrt{13}}$

u and v are orthogonal to each other, if <u, v> = 0 but here we got <u, v> = -13

It shows that u and v are not orthogonal to each other.

Theorem 3: If u and v are orthogonal vectors then for α, β any scalar we have

$\|\alpha u + \beta v||^2 = \alpha^2||u||^2 + \beta^2||v||^2$

Proof: $\|\alpha u + \beta v||^2 = << \alpha u + \beta v , \alpha u + \beta v >$

$= < \alpha u , \alpha u + \beta v > + <\beta v , \alpha u + \beta v >$ (linearity)

$= < \alpha u, \alpha u > + < \alpha u, \beta v > + < \beta v, \alpha u > + < \beta v, \beta v >$

$= \alpha^2 < u, u > + \alpha\beta < u, v > + \beta\alpha < v, u > + \beta^2 < v, u >$

$= \alpha^2||u||^2 + 2\alpha\beta < u, v > + \beta^2||v||^2$ (symmetricity)

$= \alpha^2||u||^2 + \beta^2||v||^2$ (orthogonality)

$\therefore \|\alpha u + \beta v||^2 = \alpha^2||u||^2 + \beta^2||v||^2$

Hence proved.

Properties of Orthogonality:

i. Let u, v are orthogonal vectors, then $< \alpha u, \alpha v >$ = 0, for any scalar $\alpha \in R$.

ii. If u and v are orthogonal to w then u+v is orthogonal to w.

Proof:

i. Since u and v are orthogonal to each other. ⇒ <u, v> = 0.

Multiplying α^2 both sides, $< \alpha u, \alpha v >$ = 0, for any scalar $\alpha \in R$.

ii. Given that u and v are orthogonal to w, then <u, w> = 0 and <v, w> = 0.

We have to show that <u+v, w> = 0

Consider L.H.S: <u+v, w> = <u, w> + <v, w> = 0+0 = 0 (by linearity)

Parallel and Perpendicular Vectors:

Two vectors u and v are parallel to each other if <u, v> = 1 and

If two vectors are perpendicular to each other if <u,v> = 0

Example 1: Find the vector orthogonal to both u = (-6, 4, 2) and v = (3, 1, 5).

Solution: Let x = (x_1, x_2, x_3) is orthogonal to both u and v.

x•u = $(x_1, x_2, x_3)\bullet(-6, 4, 2)$ = 0

⇒ $-6x_1 + 4x_2 + 2x_3 = 0$----------------(i)

similarly x•v = $(x_1, x_2, x_3)\bullet(3, 1, 5) = 0$

⇒ $3x_1 + x_2 + 5x_3 = 0$-----------------(ii)

Multiplying equation(ii) with 2 and then add it in equation (i), we get

$6x_2 + 12x_3 = 0 \Rightarrow x_2 = -2x_3$-------(iii)

Substituting value of x_2 in equation (ii), we get,

$x_1 = -x_3$

$$\Rightarrow x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_3 \\ -2x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$$

Hence, the vector orthogonal to both u and v is {x: x(-1, -2, 1), x∈R}

6.3 PROJECTION

Let v be a non-zero vector of a vector space V. Let W be a subspace of V. If w ∈ W is a vector such that it is closest to v, then w is called projection of v. Now decomposing an arbitrary vector x into the form x = αv + z where z ∈ $V^{\perp}$ since z ⊥ v then <v, x> = $<\alpha v, v>$ = α <v, v>.

It implies that $\alpha = \frac{<v,x>}{<v,v>}$.

The vector $\text{proj}_v^{(x)} = \frac{<v,x>}{<v,v>}\ v$ is called the orthogonal projection of x along v.

Let u be the subspace spanned by $u_1, u_2, \ldots \ldots u_n$. Then any vector v can be written as the sum of vectors in w and a vector orthogonal to W as

$$\text{proj}_{u_1,u_2,\ldots\ldots u_n}{}^{v} = \frac{v\bullet u_1}{u_1\bullet u_1}\ u_{1}+ \frac{v\bullet u_2}{u_2\bullet u_2}\ u_2 + \cdots + \frac{v\bullet u_n}{u_n\bullet u_n}\ u_n$$

$\text{proj}_{u_1,u_2,\ldots\ldots u_n}{}^{v}$ is called closest point to v in the subspace spanned by $u_1, u_2, \ldots \ldots u_n$.

The distance between the vectors v and u is $c = \frac{<v,\ u_1>}{<u_1,u_1>}$.

The point in span {u} closest to v is $v^{||u}$ =cu.

Example 1: Find the projection of v(4, 2, 1) on the vector u(5, -3, 3).

Solution: Projection of v along u = $\frac{<u,v>}{<u,u>}$

Since <u, v> = 17 and $\sqrt{<u,u>} = \sqrt{43}$

Projection = $\frac{17}{\sqrt{43}}$.

Example 2: Let a = (3, 0), b = (2, 1) find vector in span {a} that is closest to be is $b^{||a}$ and distance $||b^{\perp a}||$.

Solution: Distance $||b^{\perp a}|| = \frac{<b,a>}{<a,a>} = \frac{<(2,1),(3,0)>}{<3,0>,(3,0)>} = \frac{6}{9} = \frac{2}{3}$

$b^{||a} = \frac{<b,a>}{<a,a>}$ a= $\frac{2}{3}$*(3,0) = (2,0).

6.4 ORTHOGONAL SET OF GENERATORS

Let B = $\{v_1, v_2, \ldots\ldots, v_n\}$ be a basis of a subspace W of an inner product space V. An orthogonal Basis B' = $\{w_1, w_2, \ldots\ldots, w_n\}$ may be constructed as follows:

$$w_1 = v_1, \qquad w_1 = \text{span}\{w_1\}$$

$$w_2 = v_2 - \text{proj}_{w_1}{}^{v_2}\ ,\ \ w_2 = \text{span}\{w_1, w_2\}$$

$$\vdots$$

$$w_k = v_k - \text{proj}_{w_{k-1}}{}^{(v_k)}$$

This can be written as

$$w_1 = v_1$$

$$w_2 = v_2 - \frac{<w_1, v_2>}{<w_1, w_1>} w_1$$

$$w_3 = v_3 - \frac{<w_1, v_3>}{<w_1, w_1>} w_1 - \frac{<w_2, v_3>}{<w_2, w_2>} w_2$$

$$\vdots$$

$$w_k = v_k - \frac{<w_1, v_k>}{<w_1, w_1>} w_1 - \frac{<w_2, v_k>}{<w_2, w_2>} w_2 - \cdots - \frac{<w_{k-1}, v_k>}{<w_{k-1}, w_{k-1}>} w_{k-1}$$

The method of constructing the orthogonal vector $w_1, w_2, \ldots\ldots, w_k$ is known as the Gram-Schmidt Orthogonalization process.

Clearly, the vector $w_1, w_2, \ldots\ldots, w_k$ are linear combinations of $v_1, v_2, \ldots\ldots, v_k$. Conversely, the vectors $v_1, v_2, \ldots\ldots, v_k$ are also linear combination of $w_1, w_2, \ldots\ldots, w_k$.

Hence the basis { $w_1, w_2, \ldots\ldots, w_k$} constructed by Gram Schmidt process is an orthogonal basis of W.

Example 1: Find the orthonormal basis for subspace R^4 whose generators are $v_1 = (1, 1, 1, 1)$ $v_2 = (1,2,4,5)$, and v_3 = (1,-3,-4,-2) using Gram-Schmidt orthogonalization method.

Solution: $w_1 = v_1$ =(1,1,1,1)

$$w_2 = v_2 - \frac{<w_1,v_2>}{<w_1,w_1>} w_1$$

$$= (1, 2, 4, 5) - \frac{<(1,1,1,1),(1,2,4,5)>}{<(1,1,1,1),(1,1,,1,1)>}(1,1,1,1)$$

$$= (1, 2, 4, 5) - \frac{12}{4} (1,1,1,1)$$

$$= (1, 2, 4, 5) - (3, 3, 3, 3)$$

$$= (-2, -1, 1, 2)$$

$$w_3 = v_3 - \frac{<w_1, v_3>}{<w_1, w_1>} w_1 - \frac{<w_2, v_3>}{<w_2, w_2>} w_2$$

$$= (1,-3,-4,-2)- \frac{<(1,1,1,1),(1,-3,-4,-2)>}{<(1,1,1,1,),(1,1,1,1)>} (1,1,1,1) - \frac{<(-2,-1,1,2),(1,-3,-4,-2)>}{<(-2,-1,1,2),(-2,-1,1,2)>} (-2,-1,1,2)$$

$$= (1,-3,-4,-2)- (-2,-2,-2,-2) + \frac{7}{10} (-2,-1,1,2)$$

$$= (\frac{-1}{5}, \frac{-17}{10}, \frac{-13}{10}, \frac{7}{5})$$

Example 2: Construct an orthonormal basis of R^2 by Gram-Schmidt process S = {(3,1),(4,2)}

Solution: Let the orthonormal basis set is $\{w_1, w_2\}$

$w_1 = v_1 = (3,1)$

$$w_2 = v_2 - \frac{< w_1, v_2 >}{< w_1, w_1 >} w_1$$

$$=(4,2) - \frac{<(3,1)\bullet(4,2)>}{<(3,1)\bullet(3,1)>} \bullet (3,1)$$

$$= (4, 2) - \frac{14}{10} (3,1)$$

$$=\left(\frac{-1}{5}, \frac{-3}{5}\right)$$

6.5 ORTHOGONAL COMPLEMENT

Let W⊆ R^n be a subspace. If a vector v is orthogonal to every vector w ∈ W, we say that v is orthogonal to W. The orthogonal Complement of W is the collection of all vectors orthogonal to W. It is denoted by $W^\perp$.

i.e. $W^\perp = \{ v \in R^n : v\bullet w = 0 \text{ for all } w \in W\}$.

Theorem 4: Let W be a subset of vector space V. Prove that $W^\perp$ is a subspace of R^n.

Proof: $W^\perp$ is non-empty, since $0 \in W^\perp$ for all $w \in W^\perp, < 0, w > = 0$.

Let $w_1, w_2 \in W^\perp$.

$$<w_1 - w_2, w > = < w_1, w > + < -w_2, w > \quad \text{(linearity)}$$

$$= < w_1, w > - < w_2, w >$$

$$= 0 – 0 = 0$$

Hence we can say that $w_1, w_2 \in W^\perp$. And by the axiom of subspace we can say that $W^\perp$ is a subspace.

Theorem 5: _If $\{ w_1, w_2, \ldots, w_k\}$ forms a basis of W. then

$x \in W^\perp$ if and only if $x\bullet w_i = 0$ for all integers $1 \le i \le k$.

Proof: Let $x\bullet w_i = 0$.

Let w ∈ W, then W can be written as a linear combination of $w_1, w_2, \ldots . w_k$ as

$W = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_k w_k$.

then $x\bullet W = W = \alpha_1 x w_1 + \alpha_2 x w_2 + \cdots + \alpha_k x w_k$ (by linearity)

$= 0 + 0 + \ldots + 0 = 0$

⇨ $x \in W^\perp$.

Let $x \in W^{\perp}$. Then by definition of orthogonality $x \bullet w_i = 0$, $\forall\, w_i \in W$. Hence proved.

Theorem 6: $W^{\perp}$ is the Orthogonal Complement of W where W is a subspace of V . Then $V = W \oplus W^{\perp}$ and $W \cap W^{\perp} = \{0\}$.

Proof: We have $W \subseteq V$ and also $W^{\perp} \subseteq V$ then $W \oplus W^{\perp} \subseteq V$----(i).

Now for any $b \in V, b = b''s + b^{\perp}S$, where b” $\in$ W and $b^{\perp} \in W^{\perp}$.

$\therefore b \in W \oplus W^{\perp}$

$\Rightarrow V \subseteq W \oplus W^{\perp}$-------(ii)

From equation (i) and equation (ii), we get $V = W \oplus W^{\perp}$.

Now, $W^{\perp} = \{ v \in V: \langle v, w\rangle = 0, \forall\, w \in W\}$.

Since $W \subseteq V \Rightarrow \langle w, w\rangle = 0 \Rightarrow w = 0$.

$\therefore W \cap W^{\perp} = \{0\}$.

Hence Proved.

Example 1: Find the orthogonal Complement of W = span$\{w_1, w_2\}$, where w_1 = (3, 0, 1, 1) and w_2 = (0, 2, 5, 1).

Let $x = (x_1, x_2, x_3, x_4) \in R^4$ such that $x \bullet w_1 = x \bullet w_2 = 0$

$\Rightarrow (x_1, x_2, x_3, x_4) \bullet (3, 0, 1, 1) = 0$ and $(x_1, x_2, x_3, x_4) \bullet (0, 2, 5, 1) = 0$

$\Rightarrow 3x_1 + 0x_2 + x_3 + x_4 = 0$ and $0x_1 + 2x_2 + 5x_3 + x_4 = 0$.

We can write (x_1, x_2, x_3, x_4) in the following manner:

$x_1 = -x_3 - x_4$

$x_2 = -5x_3 - x_4$

$x_3 = 1x_3 + 0x_4$

$x_4 = 0x_3 + 1x_4$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -5 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

$\Rightarrow$ The orthogonal complement of W is $\{x_3(-1, -5, 1, 0) + x_4(-1, -1, 0, 1)$: $x_3, x_4 \in R\}$.

6.6 SUMMARY

The standard inner product of a vector v with itself gives the Euclidian length and the standard inner product of two vectors gives the angle between them. The orthogonal projection of vector w onto vector v can be assumed as shadow of w on the line spanned by v if the direction of the sun’s rays were exactly perpendicular to the line.

6.7 REFERENCE

1. Linear Algebra and Probability for Computer Science Applications, Ernest Davis, A K Peters/CRC Press (2012).

2. Linear Algebra and Its Applications, Gilbert Strang, Cengage Learning, 4th Edition (2007).

EXERCISE

Q1. Find the inner product of u and v, also show that <3u-2v, w> = 3<u, w> -2 <v, w>.

i. u=(1, -1, 2, 3) , v = (1, 0, 3, 7) and w = (2, 5, 1, 9)

ii. u = (7, 3, -9, 1), v = (2, 5, 3, 0) and w = (-1, 3, 5, 7)

iii. u = (1, 2, 3, 4), v = (2, 3, 4, 5) and w = (4, 5, 6, 7)

iv. u = (1, 9, 11, 0), v = (3, -1, 5, 7) and w = (11, 11, 5, 0)

Q2. Find the projection of vector u along vector v where u and v are,

i. u = (1, 1) and v = (1, 0)

ii. u = (0, 1) and v = $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$

iii. u = (-1, 3) and v = (3, 4)

iv. u = (-11, 10) and v = (6, 8)

Q3. Find the orthonormal basis for subspace of R^4 generated by the following:

i. (1, 2, 1, 0) and (1, 2, 3, 1)

(1, 1, 0, 0), (1, -1, 1, 1) and (-1, 0, 2, 1)

❄❄❄❄❄❄❄

7

EIGEN VECTORS

Unit Structure:

7.0 OBJECTIVES

After going to this chapter, you will be able to:

- Define discrete dynamic process
- Find eigenvalues and eigenvectors of a square matrix
- Understand Diagonalization of a matrix and its importance
- Explain Markov process, Markov Chain and Steady state
- Define Internet worm and Page rank

7.1 MODELLING DISCRETE DYNAMIC PROCESSES

A matrix equation is called a discrete dynamical system if it is in the form $\overrightarrow{x_{n+1}} = A \bullet \overrightarrow{x_n}$ or equivalently, it is $\overrightarrow{x_{n+1}} = A^{n+1} \bullet \overrightarrow{x_n}$

where A is an mxm matrix and for each integer n, $\overrightarrow{x_n}$ is an m-demensional vector.

In order to better understand the behaviour of discrete dynamical systems, we need a method of easily computing the product of matrices and vectors.

Let A = $\begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix}$, $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $v_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

(a) we are finding $A(v_1)$ and $A(v_2)$.

$$A(v_1) = \begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = v_1$$

$$A(v_2) = \begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 9 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 3v_2$$

(b) we are finding A^2v_1 and A^2v_2.

$$A^2v_1 = \begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix}^2 \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix} 3 \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$= 3 \begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$= 3\left(3 \begin{bmatrix} 3 \\ 1 \end{bmatrix}\right)$$

$$= 3^2v_2$$

Similarly we can find A^nv_1 and A^nv_2.

Based on the above procedure we can conclude that $A^nv_1 = v_1$ and $A^nv_2 = 3^nv_2$.

(c) Use the fact that $\begin{bmatrix} 7 \\ 1 \end{bmatrix} = -2v_1 + 3v_2$ to find a formula for $A^n \begin{bmatrix} 7 \\ 1 \end{bmatrix}$.

We have $\begin{bmatrix} 7 \\ 1 \end{bmatrix} = -2v_1 + 3v_2$

Multiplying both sides with A^n, we get,

$$A^n\left(\begin{bmatrix} 7 \\ 1 \end{bmatrix}\right) = A^n(-2v_1 + 3v_2)$$

$\Rightarrow \quad A^n(\begin{bmatrix}7\\1\end{bmatrix}) = A^n(-2v_1) + A^n(3v_2)$

$\Rightarrow \quad A^n(\begin{bmatrix}7\\1\end{bmatrix}) = -2A^n(v_1) + 3A^n(v_2)$

$\Rightarrow \quad A^n(\begin{bmatrix}7\\1\end{bmatrix}) = -2\, v_1 + 3(3^n v_2)$

$\Rightarrow \quad A^n(\begin{bmatrix}7\\1\end{bmatrix}) = -2\begin{bmatrix}1\\1\end{bmatrix} + 3x\; 3^n\begin{bmatrix}3\\1\end{bmatrix}$

$\Rightarrow \quad A^n(\begin{bmatrix}7\\1\end{bmatrix}) = \begin{bmatrix}-2\\-2\end{bmatrix} + \begin{bmatrix}3x3^{n+1}\\3x3^n\end{bmatrix}$

$\Rightarrow \quad A^n(\begin{bmatrix}7\\1\end{bmatrix}) = \begin{bmatrix}-2+3x3^{n+1}\\-2+3x3^n\end{bmatrix}$

$A^n(\begin{bmatrix}7\\1\end{bmatrix}) = \begin{bmatrix}-2+3^{n+2}\\-2+3^{n+1}\end{bmatrix}$ ----------(I)

It is a formula that allows us to directly compute a value by simply putting a value of n and directly getting an output.

If we want value of $A^{100}(\begin{bmatrix}7\\1\end{bmatrix})$, we can simply take n=100 in the above eqn(I) instead of multiplying A by 100 times. It allows us to compute a very large matrix multiplication very quickly and efficiently.

7.2 EIGEN VALUES AND EIGEN VECTORS

Characteristic Equation: Let A be a Square matrix, I be the unit matrix of same order that of A, and λ is a number. Then the polynomial equation det(A-λI) = 0 in the variable λ for the given square matrix A is called the characteristic equation of the matrix A.

Eigen Values: The roots of the characteristic equation det(A-λI) = 0 is called characteristics roots or eigenvalues or latent roots of the matrix A.

Eigen Vectors: An eigen vector of A is a non-zero vector v such that Av=λv , for some scalar λ. Where λ is an eigen value of A.

To find the eigenvectors of A corresponding to each eigenvalue λ, we must solve the matrix equation (A-λI)v = 0, for each eigen value λ.

Example 1: Find the characteristic equation and hence eigenvalues for A= $\begin{bmatrix}1 & -3\\-4 & 5\end{bmatrix}$.

Solution: Given A= $\begin{bmatrix}1 & -3\\-4 & 5\end{bmatrix}$.

Consider the characteristic equation as | A-λI|=0

$\Rightarrow \begin{bmatrix} 1 & -3 \\ -4 & 5 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0$

$\Rightarrow \begin{bmatrix} 1-\lambda & -3 \\ -4 & 5-\lambda \end{bmatrix} = 0$

$\Rightarrow \lambda^2$- 6 λ -7 = 0

⇒ (λ-7)(λ+1) = 0

Hence $\lambda = -1, 7$ are the eigen values for the given matrix.

Example 2: Find the characteristic equation and hence eigenvalues for A= $\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$.

Solution: Given A= $\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$.

Consider the characteristic equation as | A-λI| = 0,

$\Rightarrow \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0$

$\Rightarrow \begin{bmatrix} 1-\lambda & 2 \\ 4 & 3-\lambda \end{bmatrix} = 0$

$\Rightarrow \lambda^2$-4λ-5 = 0

⇒ (λ-5)(λ+1)=0

Hence $\lambda = -1, 5$ are the eigen values for the given matrix A.

Example 3: Find eigenvalues of matrix A= $\begin{bmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ 3 & 1 & 1 \end{bmatrix}$.

Solution: Given A= $\begin{bmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ 3 & 1 & 1 \end{bmatrix}$.

Consider the characteristic equation of A is | A-λI| = 0.

$\Rightarrow \begin{bmatrix} 1-\lambda & 1 & 3 \\ 1 & 5-\lambda & 1 \\ 3 & 1 & 1-\lambda \end{bmatrix}$=0

$\Rightarrow \lambda^3$-7λ^2+36 = 0

⇒ (λ-6)(λ-3) (λ+2) = 0

⇒ λ = -2, 3, and 6 are the eigen values for the given matrix A.

Example 4. Find eigenvalues and given vectors of A= $\begin{bmatrix} 8 & -8 & -2 \\ 4 & -3 & -2 \\ 3 & -4 & 1 \end{bmatrix}$.

Solution: Here, $A = \begin{bmatrix} 8 & -8 & -2 \\ 4 & -3 & -2 \\ 3 & -4 & 1 \end{bmatrix}$

The characteristic equations is $|A-\lambda I| = 0$.

$\Rightarrow \left|\begin{bmatrix} 8 & -8 & -2 \\ 4 & -3 & -2 \\ 3 & -4 & 1 \end{bmatrix} - \lambda\begin{bmatrix} 8 & -8 & -2 \\ 4 & -3 & -2 \\ 3 & -4 & 1 \end{bmatrix}\right| = 0$

$\Rightarrow \begin{bmatrix} 8-\lambda & -8 & -2 \\ 4 & -3-\lambda & -2 \\ 3 & -4 & 1-\lambda \end{bmatrix} = 0$

$\Rightarrow \lambda^3 - 6\lambda^2 + 11\lambda - 6 = 0$

$\Rightarrow (\lambda-1)(\lambda-2)(\lambda-3) = 0$

$\Rightarrow \lambda=1,2,3.$

Case 1: Eigen vector corresponding to eigenvalue $\lambda = 1$;

Consider $(A-1I)\, v = O$;

$\Rightarrow \begin{bmatrix} 7 & -8 & -2 \\ 4 & -4 & -2 \\ 3 & -4 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$\Rightarrow 7v_1 - 8v_2 - 2v_3 = 0$

$\Rightarrow 4v_1 - 4v_2 - 2v_3 = 0$

$\Rightarrow 3v_1 - 4v_2 = 0$

$\Rightarrow 3v_1 = 4v_2$

$\Rightarrow v_1 = 4/3 v_2$

Substituting this value of v_1 in $7v_1 - 8v_2 - 2v_3 = 0$

$\Rightarrow \frac{28}{3} v_2 - 8v_2 - 2v_3 = 0$

$\Rightarrow v_3 = \frac{2}{3} v_2$

Thus $\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} {}^4\!/_3\, v_2 \\ v_2 \\ {}^2\!/_3\, v_2 \end{bmatrix} = \frac{1v_2}{3}\begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}$

This implies that $X_1 = \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}$.

Case 2: Eigen vector corresponding to eigenvalue $\lambda = 2$;

Consider $(A-2I)v = O$;

$$\Rightarrow \begin{bmatrix} 6 & -8 & -2 \\ 4 & -5 & -2 \\ 3 & -4 & 4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Performing row operations $R_3 \rightarrow 2R_3 - R_1$ and $R_2 \rightarrow R_2 - R_1$;

$$\Rightarrow \begin{bmatrix} 6 & -8 & -2 \\ -2 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow 6v_1 - 8v_2 - 2v_3 = 0$

$\Rightarrow -2v_1 + 3v_2 = 0$

$\Rightarrow v_1 = \frac{3}{2} v_2$

Substituting the value of v_1 in $6v_1 - 8v_2 - 2v_3 = 0$,

$\Rightarrow -v_3 = \frac{v_2}{2}$

Thus $\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} v_2 \\ v_2 \\ \frac{1}{2} v_2 \end{bmatrix} = \frac{1}{2} v_2 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$

This implies that $X_2 = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$.

Case 3: Eigen vector corresponding to $\lambda = 3$:

Consider (A-3I) = O, By simplification, we get

$$\begin{bmatrix} 5 & -8 & -2 \\ 4 & -6 & -2 \\ 3 & -4 & -2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow$Performing $R_2 \longrightarrow 5R_2 - 4R_1$, and $R_3 \longrightarrow 5R_3 - 3R_1$

$$\Rightarrow \begin{bmatrix} 5 & -8 & -2 \\ 0 & 2 & -2 \\ 0 & 4 & -4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Performing $R_3 \longrightarrow R_3 - 2R_2$, we get

$$\Rightarrow \begin{bmatrix} 5 & -8 & -2 \\ 0 & 2 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow 5v_1 - 8v_2 - 2v_3 = 0 - - - - - - - - - - - - (i)$

$\Rightarrow \quad 2v_2 - 2v_3 = 0$

$\Rightarrow v_2 = v_3$ substituting in (i), we get

$\Rightarrow v_1 = 2v_3$

$$\Rightarrow \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 2\begin{bmatrix} 2v_3 \\ v_3 \\ v_3 \end{bmatrix} = v_3 \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix};$$

Hence, $X_3 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$.

Thus the eigenvalues are 1, 2, and 3. Their corresponding eigen vectors are $\begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}$, $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$ respectively.

Example 5: Find eigen values and Eigen vectors of matrix A= $\begin{bmatrix} 2 & -2 & 3 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{bmatrix}$.

Solution: Given A= $\begin{bmatrix} 2 & -2 & 3 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{bmatrix}$

Consider the characteristic equation of A is | A-λI| = 0.

$$\Rightarrow \begin{bmatrix} 2-\lambda & -2 & 3 \\ 1 & 1-\lambda & 1 \\ 1 & 3 & -1-\lambda \end{bmatrix} = 0$$

$\Rightarrow (\lambda-1)(\lambda-3)(\lambda+2) = 0$

$\Rightarrow \lambda=1,3,-2.$

Case 1: Eigen vector corresponding to eigenvalue $\lambda = 1$;

Consider (A-1I)v = O,

$$\Rightarrow \begin{bmatrix} 2-1 & -2 & 3 \\ 1 & 1-1 & 1 \\ 1 & 3 & -1-1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & -2 & 3 \\ 1 & 0 & 1 \\ 1 & 3 & -2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Performing $R_2 \longrightarrow R_2 - R_1$, and $R_3 \longrightarrow R_3 - R_1$;

$$\Rightarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 5 & -5 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Performing $R_2 \longrightarrow R_2/2$ and $R_3 \longrightarrow R_3/5$;

$$\Rightarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Performing $R_3 \longrightarrow R_3 - R_2$;

$$\Rightarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow v_1 + 2v_2 + v_3 = 0 - - - - - - - - - - - -(i)$

$\Rightarrow v_2 - v_3 = 0$

$\Rightarrow v_2 = v_3$

Substituting the value of v_2 in (i), we get

$\Rightarrow v_1 = -3v_3$

$$\Rightarrow \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -3v_3 \\ v_3 \\ v_3 \end{bmatrix} = v_3 \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}$$

Hence $X_1 = \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}$.

Case 2: Eigen vector corresponding to eigen value λ = 3

Consider (A-3I)v = O;

$$\Rightarrow \begin{bmatrix} 2-3 & -2 & 3 \\ 1 & 1-3 & 1 \\ 1 & 3 & -1-3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -1 & -2 & 3 \\ 1 & -2 & 1 \\ 1 & 3 & -4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Performing $R_2 \longrightarrow R_2 - R_1$, and $R_3 \longrightarrow R_3 + R_1$,

$$\Rightarrow \begin{bmatrix} -1 & -2 & 3 \\ 0 & 0 & -2 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Performing $R_2 \leftrightarrow R_2$,

$$\Rightarrow \begin{bmatrix} -1 & -2 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow -v_1 - 2v_2 + 3v_3 = 0 - - - - - - - - - - - -(i)$

$\Rightarrow v_2 - v_3 = 0$

$v_2 = v_3$

Substituting value of in v_2 equation (i), $-v_1 = -v_3$

$\Rightarrow v_1 = v_3$

$$\Rightarrow \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_3 \\ v_3 \\ v_3 \end{bmatrix} = v_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Hence $X_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Case 2: Eigen vector corresponding to eigenvalue λ = -2;

Consider (A+2I)v = O;

$$\Rightarrow \begin{bmatrix} 2+2 & -2 & 3 \\ 1 & 1+2 & 1 \\ 1 & 3 & -1+2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 4 & -2 & 3 \\ 1 & 3 & 1 \\ 1 & 3 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Performing $R_2 \longrightarrow 4R_2 - R_1$, and $R_3 \longrightarrow 4R_3 - R_1$;

$$\Rightarrow \begin{bmatrix} 4 & -2 & 3 \\ 0 & 14 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow 4v_1 - 2v_2 + 3v_3 = 0 - - - - - - - - - - - -$(i)

$\Rightarrow 14v_2 + v_3 = 0 - - - - - - - - - - - -$(ii)

$\Rightarrow -14v_2 = v_3$ substituting in equation (i), we get

$\Rightarrow 4v_1 - 2v_2 + 3(-14v_2) = 0$

$\Rightarrow 4v_1 - 44v_2 = 0$

$\Rightarrow 4v_1 = 44v_2$

$\Rightarrow v_1 = 11v_2$

$$\Rightarrow \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 11v_2 \\ v_2 \\ -14v_2 \end{bmatrix} = v_2 \begin{bmatrix} 11 \\ 1 \\ -14 \end{bmatrix}$$

Hence $X_3 = \begin{bmatrix} 11 \\ 1 \\ -14 \end{bmatrix}$.

Thus the eigenvalues are 1,3, and -2 and their corresponding eigenvectors are $\begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 11 \\ 1 \\ -14 \end{bmatrix}$ respectively.

Example 6: Find Eigen values and Eigen vectors of $\begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix}$.

Solution: Given A = $\begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix}$

The characteristic equation of the square matrix A is | A-λI| = 0.

i.e. $\begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix}$- λ $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ = 0

⇒ $\begin{vmatrix} 3-\lambda & -1 & 1 \\ -1 & 3-\lambda & -1 \\ 1 & -1 & 3-\lambda \end{vmatrix}$ = 0

⇒ (3 - λ)$\begin{vmatrix} 3-\lambda & -1 \\ -1 & 3-\lambda \end{vmatrix} - (-1)\begin{vmatrix} -1 & -1 \\ 1 & 3-\lambda \end{vmatrix} + 1\begin{vmatrix} -1 & 3-\lambda \\ 1 & -1 \end{vmatrix} = 0$

⇒ (3- λ)((3- λ)(3- λ) -1) + (-(3- λ) + 1) + 1(1- (3- λ))=0

⇒ (λ-2) (λ-5) (λ-2)=0

⇒ λ = 2, 2, 5 are eigenvalues of A.

Case 1: Eigen vector corresponding to eigenvalue λ = 5;

Consider (A-5I)v = O

⇒{$\begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix}$- 5 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$}$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$= O

⇒ $\begin{bmatrix} -2 & -1 & 1 \\ -1 & -2 & -1 \\ 1 & -1 & -2 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

⇒Performing row operations $R_1 \leftrightarrow R_3$, we get

⇒ $\begin{bmatrix} 1 & -1 & -2 \\ -1 & -2 & -1 \\ -2 & -1 & 1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

⇒Performing $R_2 \longrightarrow R_2 + R_1$ and $R_3 \longrightarrow R_3 + 2R_1$

⇒ $\begin{bmatrix} 1 & -1 & -2 \\ 0 & -3 & -3 \\ 0 & -3 & -3 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

⇒Performing $R_3 \longrightarrow R_3 - R_2$

⇒ $\begin{bmatrix} 1 & -1 & -2 \\ 0 & -3 & -3 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$\Rightarrow$Performing $R_2 \longrightarrow \frac{-1}{3} R_2$

$$\Rightarrow \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow V_1 - V_2 - 2V_3 = 0$

And $V_2+V_3 = 0 \Rightarrow V_2 = -V_3$

Substituting this value in $V_1 - V_2 - 2V_3 = 0$

$V_1 - V_3 = 0 \Rightarrow V_1 = V_3$

$$\Rightarrow \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_3 \\ -v_3 \\ v_3 \end{bmatrix} = v_3 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Hence, the eigenvector $X_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ is corresponding eigenvector to eigenvalue $\lambda=5$.

Case 2: Eigen vector corresponding to eigen value $\lambda = 2$

Consider $(A-2I)v = O$

$$\Rightarrow \left\{ \begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 3 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = O$$

$$\Rightarrow \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow V_1 - V_2 + V_3 = 0$

$\Rightarrow -V_1 + V_2 - V_3 = 0$

$\Rightarrow V_1 - V_2 + V_3 = 0$

We get, $V_2 = V_1 + V_3$

Now, $\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1v_1 + 0v_3 \\ 1v_1 + 1v_3 \\ 0v_1 + 1v_3 \end{bmatrix} = v_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + v_3 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

Thus , the eigenvectors are $X_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $X_3 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$.

Hence eigenvalues are 5, 2, 2 and corresponding eigenvectors are

$$X_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, X_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, X_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \text{ respectively.}$$

Properties of Eigen values:

i. The sum of the eigenvalues of a matrix is sum of the elements of principal diagonal. It is called trace of the matrix A.

$$\text{Let } A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

If λ_1, λ_2 and λ_3 are eigen values of A, then trace(A) = $\lambda_1 + \lambda_2 + \lambda_3 = a_{11} + a_{22} + a_{33}$

ii. The product of the eigenvalues of a matrix A is equal to its determinant.

iii. If λ is an eigen value of A then $\frac{1}{\lambda}$ is eigen value of A^{-1}.

iv. If $\lambda_1, \lambda_2 \dots \lambda_n$ are the eigen values of matrix A then ${\lambda_1}^m, {\lambda_2}^m, \dots\dots {\lambda_n}^m$ are eigen values of A^m.

v. If A is upper-triangular matrix of order nxn, then its eigenvalues are its diagonal elements.

vi. Eigen values of real symmetric matrix are real.

7.3 DIAGONALIZATION

If a square matrix A of order n has n linearly independent eigenvectors or n distinct eigenvalues, then there exists a matrix P of same order such that $P^{-1}AP$ is a diagonal matrix.

i.e. a square matrix A of order n is diagonalizable, iff it has n linearly independent eigen vectors.

Let $\lambda_1, \lambda_2, \dots, \lambda_n$ are the distinct eigenvalues of a matrix A of order n and the corresponding eigen vectors are $X_1, X_2, \dots, X_n$. Then a square matrix P can be formed with these eigen vectors as

$$P = [X_1\ X_2 \dots X_n].$$

Now AP = A$[X_1\ X_2 \dots X_n]$ = $[\lambda_1 X_1\ \ \lambda_2 X_2 \dots\ \lambda_n X_n]$

For n = 3, AP = A$[X_1\ X_2\ X_3]$ = $[AX_1\ \ AX_2\ \ AX_3]$ = $[\lambda_1 X_1\ \ \lambda_2 X_2\ \ \lambda_3 X_3]$

$$= \begin{bmatrix} \lambda_1 x_1 & \lambda_2 x_2 & \lambda_3 x_3 \\ \lambda_1 y_1 & \lambda_2 y_2 & \lambda_3 y_3 \\ \lambda_1 z_1 & \lambda_2 z_2 & \lambda_3 z_3 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix} * \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = PD,$$

where D is the diagonal matrix.

So $P^{-1}AP = P^{-1}PD = D$.

The matrix P which diagonalizes A is called the transforming matrix or modal matrix of A. The resulting diagonal matrix D is known as spectral matrix of A. D has the eigenvalues of A as its elements.

7.3.1 Similar Matrix:

A square matrix B of order n is called similar to a square matrix A of same order if

$$B = P^{-1}AP \text{ for some non-singular matrix P of order n.}$$

Since matrix B is similar to matrix A, B has same eigenvalues as A. If X is an eigenvector of A, then $y = P^{-1}X$ is an eigen vector of B corresponding to same eigen values.

7.3.2 Calculation of powers of a matrix:

Let matrix P diagonalizes matrix A, i.e. $D = P^{-1}AP$

Then $D^2 = (P^{-1}AP)(P^{-1}AP) = P^{-1}APP^{-1}AP = P^{-1}A^2P$

[since $PP^{-1} = I$]

Again $D^3 = (P^{-1}A^2P)(P^{-1}AP) = P^{-1}A^2P\,P^{-1}AP = P^{-1}A^3P$

Similarly, $D^n = P^{-1}A^nP$; pre-multiplying by P and post-multiplying by P^{-1}, we get

$$PD^nP^{-1} = PP^{-1}A^nPP^{-1} = A^n .$$

7.3.3 Diagonalization of the Fibonacci Matrix

Fibonacci considered the following problem (breeding rabbits):

We breed rabbits, starting with one pair of rabbits. Each pair of rabbits produces one pair of offspring in every month. After one month, the offspring is adult and ready for reproduction. After neglecting all kinds of effects (as death) and always considering pairs of rabbits, we get the number of rabbits increase quite rapidly.

Let rabbit vector $\vec{r} = \binom{j}{a} \in R^2$, where j and a denote the number of juvenile pairs and number of adult pairs respectively.

Since, $j_{n+1} = a_n$

$a_{n+1} = j_n + a_n$

In vector notation, $\begin{pmatrix} j_{n+1} \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} j_n \\ a_n \end{pmatrix}$

Or, $\vec{r}_{n+1} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \vec{r}_n$

The transition matrix of this dynamical system is A = $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$

The initial condition is $\vec{r}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, that means there is one pair of juvenile rabbits, no adult rabbits.

That means the dynamical system in the equations can be summarized as

$$\vec{r}_{n+1} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \vec{r}_n \text{ , } \vec{r}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

The solution of the above equation will be in the form of $\vec{r}_n = A^n \vec{r}_0$

Or, $\begin{pmatrix} j_n \\ a_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Analysis of the problem:

Now we calculate first few rabbit vectors :

N	0	1	2	3	4	5	6	7	8
j_n	1	0	1	1	2	3	5	8	13
a_n	0	1	1	2	3	5	8	13	21

The table shows that after 5 months, there are 3 juvenile pairs and 5 adult pairs of rabbits.

The sequence a_n = (0, 1, 1, 2, 3, 5, 8, …) is the famous Fibonacci sequence.

For finding the eigenvalue, rewrite the vector equation as follows:

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} j \\ a \end{pmatrix} = \lambda \begin{pmatrix} j \\ a \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} j \\ a \end{pmatrix} = \lambda \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} j \\ a \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} j \\ a \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} j \\ a \end{pmatrix}$$

$$\begin{pmatrix} \lambda & -1 \\ -1 & \lambda - 1 \end{pmatrix} \begin{pmatrix} j \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The rabbit vector $\begin{pmatrix} j \\ a \end{pmatrix}$ has to be non-zero, so for the solution of the above matrix equation the coefficient matrix must be zero. The matrix is singular if $\det\begin{pmatrix} \lambda & -1 \\ -1 & \lambda - 1 \end{pmatrix} = 0$.

$\det\begin{pmatrix} \lambda & -1 \\ -1 & \lambda - 1 \end{pmatrix} = \lambda(\lambda - 1) - 1 = 0$

$\lambda(\lambda-1)$-1 = 0 is called characteristic equation of the matrix A = $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ and its root is known as eigen values of A.

Solving the characteristic equation, we get the two eigen values are

$\lambda_1 = \frac{1+\sqrt{5}}{2}$ and $\lambda_2 = \frac{1-\sqrt{5}}{2}$.

Eigen vector corresponding to eigenvalue $\lambda_1 = \frac{1+\sqrt{5}}{2}$:

$$\begin{pmatrix} \frac{1+\sqrt{5}}{2} & -1 \\ -1 & \frac{1+\sqrt{5}}{2}-1 \end{pmatrix}\begin{pmatrix} j \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

After row reducing the coefficient matrix, we get $\begin{pmatrix} \frac{1+\sqrt{5}}{2} & -1 \\ -1 & \frac{1+\sqrt{5}}{2}-1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & \frac{1-\sqrt{5}}{2} \\ 0 & 0 \end{pmatrix}$

For the non-trivial solution of the above equation is $\begin{pmatrix} j \\ a \end{pmatrix} = \begin{pmatrix} \frac{(\sqrt{5}-1)a}{2} \\ a \end{pmatrix}$

Thus the eigen vector is $\vec{v}_1 = \begin{pmatrix} \sqrt{5}-1 \\ 2 \end{pmatrix}$ corresponding to eigen value eigen value $\lambda_1 = \frac{1+\sqrt{5}}{2}$.

Now we calculate eigen vector corresponding to eigen vector $\lambda_2 = \frac{1-\sqrt{5}}{2}$:

Putting the value of λ_2 in $\begin{pmatrix} \lambda & -1 \\ -1 & \lambda-1 \end{pmatrix}$, we get $\begin{pmatrix} \frac{1-\sqrt{5}}{2} & -1 \\ -1 & \frac{1-\sqrt{5}}{2}-1 \end{pmatrix}\begin{pmatrix} j \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

After row reducing the coefficient matrix, we get $\begin{pmatrix} \frac{1-\sqrt{5}}{2} & -1 \\ -1 & \frac{1-\sqrt{5}}{2}-1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & \frac{1+\sqrt{5}}{2} \\ 0 & 0 \end{pmatrix}$

For the non-trivial solution of the above equation is $\begin{pmatrix} j \\ a \end{pmatrix} = \begin{pmatrix} \frac{-(1+\sqrt{5})a}{2} \\ a \end{pmatrix}$.

Thus the eigenvector is $\vec{v}_2 = \begin{pmatrix} 1+\sqrt{5} \\ -2 \end{pmatrix}$ corresponding to eigen value eigen value $\lambda_2 = \frac{1-\sqrt{5}}{2}$.

Example 1: Find a matrix P which transforms the matrix A = $\begin{bmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ 3 & 1 & 1 \end{bmatrix}$ to diagonal form.

Hence calculate A^4 .

Solution: The characteristic equation is | A – λI | = 0

⇨ $\begin{vmatrix} 1-\lambda & 1 & 3 \\ 1 & 5-\lambda & 1 \\ 3 & 1 & 1-\lambda \end{vmatrix}$ = 0

⇨ λ^3 - $7\lambda^2$ + 36 = 0

⇨ (λ+2) (λ-3) (λ-6) = 0

⇨ Eigen values of A are λ = -2, 3 and 6.

Now eigen vector corresponding to λ = -2 can be found by solving

$\begin{bmatrix} 1+2 & 1 & 3 \\ 1 & 5+2 & 1 \\ 3 & 1 & 1+2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ = 0 i.e.

3 x + y + 3z = 0, x + 7y + z = 0 , 3x + y + 3z = 0.

We get $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ = k $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.

Similarly, eigenvectors corresponding to λ = 3 and λ = 6 are arbitrary non-zero multiples of the vectors [1, -1, 1] and [1, 2, 1].

Hence the transforming matrix P = $\begin{bmatrix} -1 & 1 & 1 \\ 0 & -1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$.

Now find P^{-1} using adjoint method :
A_{11} = -3 , A_{12} = 2 , A_{13} = 1, A_{21} = 0 , A_{22} = -2 , A_{23} = 2, A_{31} = 3, A_{32} = 2, A_{33} =1 and |P| = 6.

Hence P^{-1} = 1/6 $\begin{bmatrix} -3 & 0 & 3 \\ 2 & -2 & 2 \\ 1 & 2 & 1 \end{bmatrix}$.

Thus, D = P^{-1}AP = $\begin{bmatrix} -2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix}$

Now A^4 = $\begin{bmatrix} -1 & 1 & 1 \\ 0 & -1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 16 & 0 & 0 \\ 0 & 81 & 0 \\ 0 & 0 & 1296 \end{bmatrix}$1/6 $\begin{bmatrix} -3 & 0 & 3 \\ 2 & -2 & 2 \\ 1 & 2 & 1 \end{bmatrix}$

$= \begin{bmatrix} 251 & 485 & 235 \\ 485 & 1051 & 485 \\ 235 & 485 & 251 \end{bmatrix}$

7.4 COORDINATE REPRESENTATION IN TERMS OF EIGEN VECTORS

Let λ_1, λ_2,…., λ_n are the eigen values of a matrix A of order n and the corresponding eigen vectors are $X_1, X_2,\ldots, X_n$ which are columns of P. Let $u^{(t)}$ be the coordinate representation of $x^{(t)}$ in terms of eigenvectors. The equation $x^{(t)} = A^t . x^{(0)}$ gives rise to

$$\left[u^{(t)}\right] = \begin{bmatrix} \lambda_1{}^t & & \\ & \lambda_2{}^t & \\ & & \lambda_3{}^t \end{bmatrix} \left[u^{(0)}\right]$$

As the power increases and if $|\lambda_i| > |\lambda_j|$ for all j, then $\lambda_i{}^t$ will dominate.

7.5 THE INTERNET WORM

An Internet worm is a program that exploits flaws in utility programs in systems. The flaws allow the program to break into those machines and copy itself, thus infecting those systems.

It spread itself without human intervention by using a scanning strategy to find vulnerable hosts to infect. Some of the famous examples of code red, SQL Stammer, and Blalter. It performs self-replication by sending copies of their codes in network packets and ensuring the codes are executed by the computers that receive them. Meanwhile, when computers on network become a victim of its infection, it spreads further copies of the worm by exploiting low level software defects.

The following are the activities of worms:

i. **Infection:** By injecting new code and new control flow edges into the program. Worms gain control of the execution of a remote program.

ii. **Spreading:** Worms typically replicate itself to infect other computers.

iii. **Hiding:** Worms use the following techniques to avoid being detected on internet.

Traffic shopping, Polymorphism, and finger printing.

In order to defend against future worms, it is important to understand how worms propagate and how different scanning strategies affect worm propagation dynamies.

An efficient and reliable vigilante system for worm containment was developed using Markov chain. Markov chain is a mathematical system that describes transitions from one state to another, between a finite or countable number of possible states. The Markov chain model is developed for uniform scanning worms, specifically for scanning worms, we are able to provide condition that determines whether the worm spread would eventually stop and obtain the distribution of the total number of infected hosts.

Modelling the Worm: Worm population represented by a vector X = $[x_1, y_1, x_2, y_2, x_3, y_3,]$ for i=1,2,3 is the expected number of mortal worms at computer x_i and y_i is the expected number of immortal worms at computer i.

For t = 0, 1, 2, …. Let $x^{(t)} = \{x_1{}^{(t)}, x_2{}^{(t)}, y_1{}^{(t)}, y_2{}^{(t)}, x_3{}^{(t)}, y_3{}^{(t)}\}$, any mortal worm on computer 1 is a child of computer 2 or 3.

Therefore, the expected number of mortal worms at computer 1 after t+1 iterations is

$$x_1{}^{(t+1)} = \frac{1}{10}x_2{}^{(t)} + \frac{1}{10}y_2{}^{(t)} + \frac{1}{10}x_3{}^{(t)} + \frac{1}{10}y_3{}^{(t)}$$

With probability $\frac{1}{7}$, a mortal worm at computer 1 becomes immortal. The previously immortal worms stay immortal. Therefore, $y_1{}^{(t+1)} = \frac{1}{7}x_1{}^{(t)} + y_1{}^{(t)}$

Then we get a matrix A such that

$$A=\begin{bmatrix} 0 & 0 & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} \\ \frac{1}{7} & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{10} & \frac{1}{10} & 0 & 0 & \frac{1}{10} & \frac{1}{10} \\ 0 & 0 & \frac{1}{7} & 1 & 0 & 0 \\ \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{7} & 1 \end{bmatrix}$$

The matrix has linearly independent vectors and its largest eigenvalues is about 1034.

7.6 EXISTENCE OF EIGEN VALUES

If A is an n*n matrix with entries in C, Then det(A - λ I) is a polynomial of degree n in λ with coefficients in C. By the corollary of fundamental theorem of algebra, it has n roots. This gives the existence of eigenvalues. Let V ≠ {0} be a finite dimensional vector space over C, and let T ∈ L(V, V). Then, T has at least one eigenvalue.

7.7 MARKOV CHAINS

Stochastic Process: Stochastic process is a process that involves a variable changing at a random rate through time. There are various types of stochastic process such as random walks, Markov chains and Bernoulli processes.

Probability vectors: A row vector v = $(v_1, v_2, \dots, v_n)$ is called a probability vector if $v_1, v_2, \dots, v_n$ are non-negative and their sum is equal to 1.

Example: v = $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, 0)$

Markov Property: A Markov property or memoryless property, when the future and past states are given, the future states of the process depend only on present state and not at all the past states.

Markov Process: A random process with the Markov property is called Markov process.

Markov Chain: A Markov chain is a Markov process with discrete time and discrete state space. Markov chain is a mathematical model that describes transitions from one state to another according to certain probabilistic rules. It is a stochastic process in which possible future states are fixed. In other words, the probability of transitioning to any particular state is dependent only on the current state and time elapsed.

Markov chain is denoted by X = $(X_n)_{n \in N} = (X_0, X_1, X_2, \dots\dots)$

7.7.1 Transition Matrix:

A Markov chain $\{X\}$ at time t can be represented as a matrix. This matrix contains information on the probability of transitioning between states, so the matrix is known as transition matrix. It is denoted by p_t. The $(i,j)^{th}$ element of the matrix p_t is given by

$(p_t)_{i,j} = p(x_{t+1} = j/x_t = i)$

This means each row of the matrix is a probability vector and the sum of its entries is 1.

Transition matrices have the property that the product of subsequent ones describes a transition along the time interval spanned by the transition matrices.

Let us assume that we have a finite number N of possible states in E such that E = $\{e_1, e_2, \dots, e_N\}$.

The initial probability distribution can be described as a row vector q_0 of size N such that $(q_0)_i = q_0(e_i) = P(X_0 = e_i)$

⇨ $P_{ij} = P(e_i, e_j) = P(x_{n+1} = e_j / x_n = e_i)$

⇨ $(q_n)_i = q_n(e_i) = P(x_n = e_i)$

⇨ $q_{n+1} = q_n P, q_{n+2} = q_{n+1} P = (q_n P) P = q_n P^2$

⇨ $q_{n+m} = q_n P^2$

That means the probability vector after n repetitions of the experiment is $q_0 P^n$.

⇨ The row vector describing probability describing at time step n+1

⇨ Row vector describing probability distribution at times of X transitions.

Properties of transition Matrix:

1. It is square, since all possible states must be used behaviors and as columns.

2. All entries are between 0 and 1, because all entries represent probability.

3. The sum of the entries in any row must be 1, since the numbers in the row give the probability of changing from the state at the left to one of states indicated across the top.

7.7.2 Graphical Representation:

The finite state space Markov chain can be represented as a directed graph such that each node in the graph is a state. For all pairs of state (e_i, e_j) there exists an edge if $P(e_i, e_j) > 0$ and the value of the edge is same probability $P(e_i, e_j)$.

Example 1: Consider the daily behavior of student of SYCS towards visit of college library for each day, there are 3 possible states: The student does not visit the library this day (N), the student visits library but does not issue any book (V) and the student visits library and takes at least one book (R) so, we have the following state space E = {N, V, R}.

Assume that at the first day this student has 70% chance to only visit library and 30% chance to visit library and to take at least one book for some the vector describing the initial probability distribution (n=0) is that $q_0 = (0.0, 0.7, 0.3)$.

Now assume that the following probabilities have been observed:

i. When the student does not visit library a day, he has 25% chance of visiting the next day, 50% chance to only visit and 25% chance to visit and to issue at least one book.

ii. When the student visits library without issuing any book a day. He has 60% chance to visit again without issuing the next day and 40% to visit and issue.

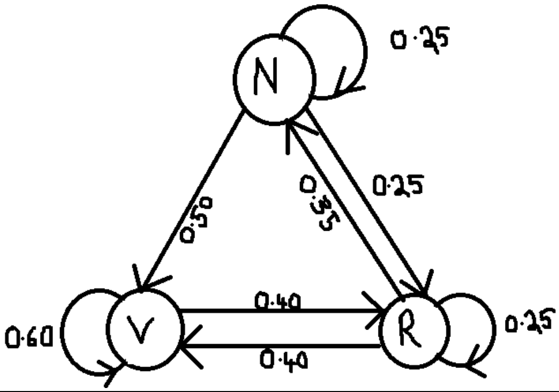
iii. When the student visits and issues a book on a day, he has 35% chance of not visiting the next day, 40% chance to only visit and 25% to visit and issue a book again.

Thus we have the transition matrix P = $\begin{array}{c} \\ \dot{N} \\ V \\ R \end{array} \begin{bmatrix} \overbrace{\begin{array}{ccc} N & V & R \end{array}} \\ \begin{array}{ccc} 0.25 & 0.5 & 0.25 \\ 0.00 & 0.60 & 0.40 \\ 0.35 & 0.40 & 0.15 \end{array} \end{bmatrix}$

The probability of each state for the second day (n=1)

$$q_1 = q_0 P = (0.0, 0.7, 0.3)\begin{bmatrix} 0.25 & 0.5 & 0.25 \\ 0.00 & 0.60 & 0.40 \\ 0.35 & 0.40 & 0.15 \end{bmatrix}$$

Finally, the probabilistic dynamic of this Markov chain can be graphically represented as follows:



7.7.3 Regular transition Matrices:

Markov chain is used to find long range predictions. It is not possible to make long range predictions with all transition matrices, but for a large set of transition matrices, predictions are possible with regular transition matrices.

A transition matrix is regular if some power of the matrix contains all positive entries. A Markov chain is a regular Markov chain if its transition matrix is regular.

This matrix L gives the long range trend of the Markov chain. It can be found by solving a system of linear equations.

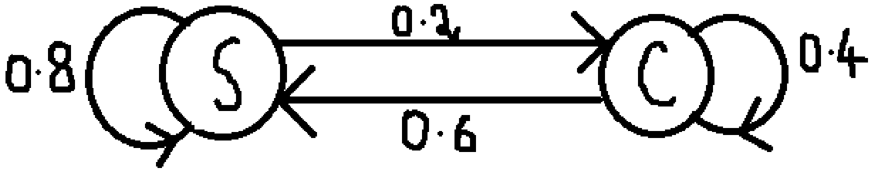
7.7.4 Steady state:(solution set):

(Equilibrium Matrix): A probability matrix which is the solution to LP=L is called equilibrium Matrix.

Absorbing Markov Chain: A state S_i of a Markov chain is called absorbing if it is not possible to leave it. A Markov chain is absorbing if it has at least one absorbing state.

Example 2: After close analysing the weather for several years, a meteorologist concludes: The chance of a day after a sunny day is sunny 80% and cloudy 20% of the time. The chance of a day after a cloudy day is sunny 60% and cloudy 40% of time. Find the long range trend.

Solution: The diagram of the Markov chain for this process having two states sunny(S) and cloudy(C) is



The transition matrix P= $\begin{array}{c c} & \begin{array}{cc} S & C \end{array} \\ \begin{array}{c} S \\ C \end{array} & \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} \end{array}$

To find long term probabilities, we have to solve LP=L where L = $[v_1 \quad v_2]$.

⇨ $[v_1 \quad v_2]\begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} = [v_1 \quad v_2]$

⇨ $0.8v_1 + 0.6v_2 = v_1$ and $0.2v_1 + 0.4v_2 = v_2$

⇨ $-0.2v_1 + 0.6v_2 = 0$ and $0.2v_1 - 0.6v_2 = 0$

⇨ Both the equations are same

⇨ $0.2v_1 = 0.6v_2$

⇨ $v_1 = 3v_2$

But we have $v_1 + v_2 = 1$ (probability vector)

Solving these equations, we have $v_1 = \frac{3}{4}$ and $v_2 = \frac{1}{4}$

Hence L= $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} \\ \frac{1}{4} \end{bmatrix}$

This vector L= $\begin{bmatrix} \frac{3}{4} \\ \frac{1}{4} \end{bmatrix}$ is a long term, the probability that the process will be in state 1 is $\frac{3}{4}$ and the probability that the process will be in state 2 is $\frac{1}{4}$.

Example 3**:** Assume that a man's profession can be classified as professional, skilled labourer or unskilled labourer. Assume that, of the sons of professional men, 80% are professional, 10% are skilled labourers and 10% are unskilled labourers. In the case of sons of skilled labourers 60% are skilled labourers, 20% are professionals and 20% are unskilled. Finally, in the case of unskilled labourers, 50% of the sons are unskilled labourers, and 25% each are in the other two categories. Assume that every man has at least one son, and form a Markov chain by following the profession of a randomly chosen son of a given family trough several generations. Form the transition matrix and find probability of their long run behaviour.

Solution: Let P, S and U denote the three states as professional, skilled labourer and unskilled labourer. According to the given information, Transition matrix P is

$$P = \begin{array}{c} \\ P \\ S \\ U \end{array} \begin{bmatrix} \overbrace{\begin{matrix} 0.8 & 0.1 & 0.1 \end{matrix}}^{P \quad S \quad U} \\ \begin{matrix} 0.2 & 0.6 & 0.2 \end{matrix} \\ \begin{matrix} 0.25 & 0.25 & 0.5 \end{matrix} \end{bmatrix}$$

Let L = (x_1, x_2, x_3) be probability vector. Then long term behaviour can be found by solving L*P = L.

⇨ $$(x_1, x_2, x_3) \begin{array}{c} \\ P \\ S \\ U \end{array} \begin{bmatrix} \overbrace{\begin{matrix} 0.8 & 0.1 & 0.1 \end{matrix}}^{P \quad S \quad U} \\ \begin{matrix} 0.2 & 0.6 & 0.2 \end{matrix} \\ \begin{matrix} 0.25 & 0.25 & 0.5 \end{matrix} \end{bmatrix} = (x_1, x_2, x_3)$$

⇨ Then $0.8x_1 + 0.2x_2 + 0.25x_3 = x_1$

⇨ $-0.2x_1 + 0.2x_2 + 0.25x_3 = 0$----------------(i)

⇨ $0.1x_1 + 0.2x_2 + 0.25x_3 = x_2$

⇨ $0.1x_1 - 0.4x_2 + 0.25x_3 = 0$----------------------(ii)

And $0.1x_1 + 0.2x_2 + 0.5x_3 = x_3$

⇨ $0.1x_1 + 0.2x_2 - 0.5x_3 = 0$--------------------------(iii)

Equation (iii) is the sum of equation (i) and equation (ii).

From equation (i), we get

$5x_3 = 4x_1 - 4x_2$---------------------------------(iv)

And from (iii), we get

$5x_3 = x_1 + 2x_2$---------------------------(v)

Solving(iv)and(v), weget,

$x_1 = 2x_2$ and $x_3 = \frac{4}{5}x_2$

Since L is the probability vector, hence, $x_1 + x_2 + x_3 = 1$

⇨ $2x_2 + x_2 + \frac{4}{5}x_2 = 1$

⇨ $x_2 = \frac{5}{19}$

⇨ Now we have L= $\begin{bmatrix} \frac{10}{19} \\ \frac{5}{19} \\ \frac{4}{19} \end{bmatrix}$.

7.8 MODELLING A WEB SURFER: PAGERANK

A search query with Google's search engine usually returns a very large number of pages. Google assigns a number to each individual webpage based on the link structure of the web, expressing its importance. This number is known as the page rank and is computed via the page rank algorithm. It has applications in search, browsing and traffic estimation.

7.8.1 Page Rank algorithm as a Markov Process:

We describe page rank algorithm as a Markov process, web page as state of Markov chain, Link structure of web as transitions probability matrix of Markov chains. It mainly focus on how to relate the eigenvalues and eigen vector of Google matrix to page rank values to guarantee that there is a single stationary distribution vector to which the page rank algorithm converges and efficiently compute the page rank for large sets of web pages.

7.8.2 Basic Page Rank Algorithm Model:

A webpage U's page rank is calculated base on how many other webpages backlink into U. The page rank of U is the sum of the page ranks of each webpage v_i that back links to U divided by the number of webpages to which v_i links. That means if webpage U is linked to only low page ranks web pages, it may not get more importance. Moreover If U is linked by a webpage v_j with a high page rank, but v_j links to many other pages, U should not receive the full weight of v_j's page rank.

Let U=web page

F_u=Forward links from U

B_u =Back links into U

C=normalization factor so that the total rank of all web pages is constant.

Then page rank (by simple rankin)=R(U)=C$\sum_{v \in B_u} \frac{R(V)}{N_V}$ where C<1

7.8.3 Random web surfer Model:

Page rank can also be defined as the model of a random web surfer navigating the internet. That means the model states that the page rank models the behavior of someone who keeps clicking on successive links at random.

Consider a simple link structure of web pages:

We can represent this structure using NXN adjacency matrix A, where $A_{ij} = 1$ if there is a link from webpage i to webpage j, and 0 otherwise.

Let N=total number of webpages in the web.

$\pi^T = 1XN\ page\ rank\ row\ vector\ (stationary\ vector)$

H=NXN row normalized adjacency matrix(or) Transition probability matrix

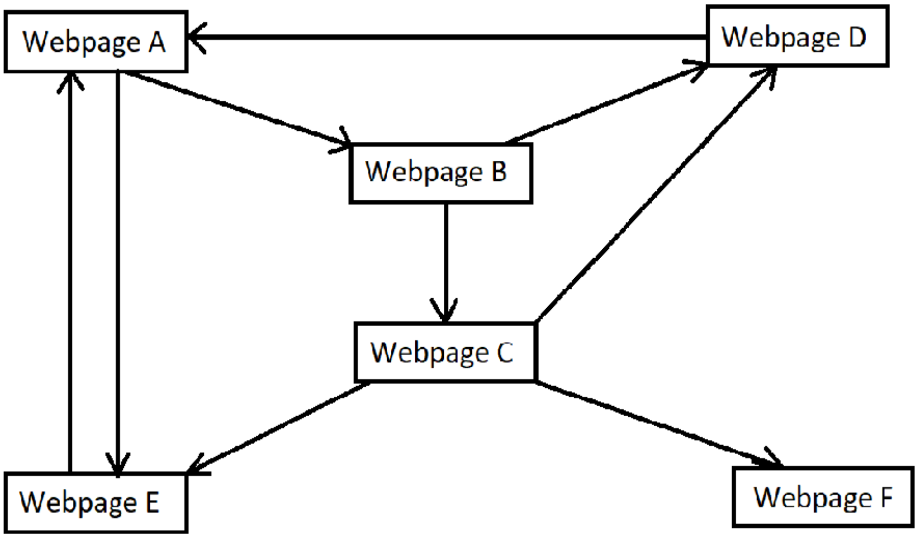
Thus we can describe the page rank vector at the k^{th} iteration as,

$\pi^{kT} = \pi^{(k-1)T}H$

To build a transition probability matrix $H_i = \frac{A_i}{\sum_{k=1}^{N} A_{ik}}$

So that each row A_i of A is divided by its row sum.

Consider the following diagram that shows the link of web pages A,B,C,D ,E and F



The 6X6 adjacency matrix for the above link structure is A = $\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Transition probability matrix H is H= $\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

But the matrix H is not stochastic due to dangling node F which has no outgoing links. It affects the model because it is not clear where its weight should be distributed. To overcome this type of problem, we assign artificial link to dangling node.

Therefore, we define the stochastic S as,

S= H+$\frac{a * e^T}{N}$

Where a = NX1 column vector such that $a_i \quad = 1\ if, \sum_{k=1}^{N} H_{ik} = 0$

= 0, otherwise.

E=NX1 column vector of one's

For the above example:S= $\begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$

It makes sure that the surfer's random walk process does not get stuck and the web pages are the states of the Markov chain.

7.8.4 Google matrix: The above matrix S has a unique stationary distribution vector π^T , if S is irreducible as well as stochastic. A matrix is irreducible if and only if its graph is strongly connected. So, we define the irreducible row stochastic matrix G as

G= $\alpha S + (1-\alpha)E;\ 0 \le \alpha \le 1\ and\ E = \frac{eXe^T}{N}$

G is the Google matrix defined as
$\pi^{kT} = \pi^{(k-1)T}G$ as the new iterative method for page rank.

For this above example:

G= $\begin{bmatrix} \frac{1}{40} & \frac{9}{20} & \frac{1}{40} & \frac{1}{40} & \frac{9}{20} & \frac{1}{40} \\ \frac{1}{40} & \frac{1}{40} & \frac{9}{20} & \frac{9}{20} & \frac{1}{40} & \frac{1}{40} \\ \frac{1}{40} & \frac{1}{40} & \frac{1}{40} & \frac{77}{250} & \frac{77}{250} & \frac{77}{250} \\ \frac{7}{8} & \frac{1}{40} & \frac{1}{40} & \frac{1}{40} & \frac{1}{40} & \frac{1}{40} \\ \frac{7}{8} & \frac{1}{40} & \frac{1}{40} & \frac{1}{40} & \frac{1}{40} & \frac{1}{40} \\ \frac{83}{500} & \frac{83}{500} & \frac{83}{500} & \frac{83}{500} & \frac{83}{500} & \frac{83}{500} \end{bmatrix}$

<u>**The power method:**</u>

The Google matrix G is currently of size max than eight billion webpages. So the Eigen value competition not so easy.

We iterate using the Google matrix G by writing $\pi^{kT} = \pi^{(k-1)T}G$

When dealing with large data sets, it is difficult to form a matrix G. It is more efficient to compute page rank vector using the power method, where we iterate using the sparse matrix H by rewriting the above equation as,

$$\pi^{kT} = \pi^{(k-1)T}G$$

$$= \pi^{(k-1)T}(\alpha S + (1-\alpha)E$$

$$= \pi^{(k-1)T}(\alpha S + (1-\alpha)(\frac{eXe^T}{N})$$

$$= \alpha\pi^{(k-1)T}\mathrm{S} + (1-\alpha)\,\pi^{(k-1)T}(\frac{eXe^T}{N})$$

$$= \alpha\pi^{(k-1)T}\mathrm{S} + (1-\alpha)\,\frac{e^T}{N}$$

$$= \alpha\pi^{(k-1)T}(H + \frac{aXe^T}{N}) + (1-\alpha)\,\frac{e^T}{N}$$

$$= \alpha\pi^{(k-1)T}H + (\,\alpha\pi^{(k-1)T}\mathrm{a} + (1-\alpha)\,\frac{e^T}{N}$$

Since $\pi^{(k-1)T}$ is a probability vector and thus $\pi^{(k-1)T}$e = 1.

The size of the Markov matrix makes storage issues non-trivial. For modern web structure for which the transition probability matrix H can be stored in main memory, compression of the data is not essential. In order to compute the page rank vector, the page rank power method requires vector matrix multiplication of $\pi^{(k-1)T}H$ at each iteration k.

Hence we can say page rank is a global ranking of all web pages, regardless of their content based solely on their location in the web's link structure. Using page rank, we are able to order search results so that more important and control webpages are given preference.

7.9 SUMMARY

Any scalar λ and vector v that satisfies the relationship Av = λv are called an eigenvalue and an eigenvector respectively of the square matrix A. Eigenvalues and eigenvectors for a linear transformation T: V → V are determined by locating the eigenvalues and eigenvectors of any matrix representation for T; the eigenvectors of the matrix are coordinate representations of the eigenvector of T. An n*n matrix is diagonalizable if and only if it has n linearly independent eigenvectors.

7.10 REFERENCES

1. Linear Algebra and Probability for Computer Science Applications, Ernest Davis, A K Peters/CRC Press (2012).

2. Linear Algebra and Its Applications, Gilbert Strang, Cengage Learning, 4th Edition (2007).

Exercise

Q1. Find Eigen values and Eigen vectors for the following

i. $\begin{bmatrix} 2 & -1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$

ii. $\begin{bmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{bmatrix}$

iii. $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$

iv. $\begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$

v. $\begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}$

Q2: Check whether the following matrices are diagonalizable or not, if yes, diagonalize them:

i. $\begin{bmatrix} -1 & 2 & 2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$

ii. $\begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{bmatrix}$

iii. $\begin{bmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{bmatrix}$

iv. $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ -4 & 4 & 3 \end{bmatrix}$

v. $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

vi. $\begin{bmatrix} 3 & 10 & 5 \\ -2 & -3 & -4 \\ 3 & 5 & 7 \end{bmatrix}$

Q3: Find the long term probability vector for the following Markov Process:

i. In the dark ages, Harward, Dard mouth and Yale admitted only male students. Assume that , at that time 80% of the sons of Harward men went to Harward and rest went to Yale, 40% of the sons of Yale men went Yale, and the rest split evenly between Harward and Dard mouth; and of the sons of Dard mouth men, 70% went to Dard mouth, 20% went to Harward, and 10% to Yale. Formulate Markov chain and find probability of their long term behavior.

ii. A salesman's territory consists of 3 cities A,B and C . He never sells in the same city on successive days. If he sells in city A, then the next day he sell in B. However if he sells in either B or C, then the next day he is twice likely to sell in city A as in the other city. In long run, how often does he sell in each of the cities?

iii. Two boys u_1 *and* u_2 and two girls g_1 *and* g_2 are throwing a ball each other.Each boy throws the ball to the other boy with probability $\frac{1}{2}$ and each to the girl with probability $\frac{1}{4}$. On the other hand , each throws the ball to each boy with probability $\frac{1}{2}$ and never to the other girls. In the long run, how often does each receive the ball.

A man walks along a four-block stretch of part-Avenue. If he is at corner 1,2,or3 then he walks to the left or right with equal probability. He continues until he reaches corner 4, which is a restaurant or corner 0, which is his home. If he reaches either home or restaurant, he stays there. Formulate the transition matrix for states 0,1,2,3 and 4 as a Markov chain.

❄❄❄❄❄❄❄