DISTANCE LEARNING CENTRE

Ahmadu Bello University
Zaria, Nigeria

MATH 208:
Linear Algebra II

Course Material

Programme Title: BSc. Computer Science

Published and Printed by

Ahmadu Bello University Press Limited, Zaria,
Kaduna State, Nigeria.
Tel: 08065949711
E-mail: abupress2013@gmail.com,
abupress2020@yahoo.com Website: www.abupress.com.ng

Course Writers/
Development Team

Editor
Prof. M.I Sule

Course Materials Development Overseer
Dr. Usman Abubakar Zaria

Subject Matter Expert
MUHAMMAD, Abubakar Alhassan

Subject Matter Reviewer
Dr. Abdulssamad Tanko Imam

Language Reviewer
Ene Ojokojo

Instructional Designers/Graphics
Emmanuel Ekoja / Ibrahim Otukoya

Course Coordinator
Emmanuel Ekoja

ODL Expert
Dr. Abdulkarim Muhammad

ACKNOWLEDGEMENTS

We acknowledge the use of the Courseware of the National Open University of Nigeria (NOUN) as the primary resource. Internal Reviewers in the Ahmadu Bello University have also been duly listed.

Contents

Course Study Guide

Course Information

Course Code:MATH208

Course Title: Linear Algebra II

Credit Units: 3 Credit Units

Year of Study: 200 Level

Semester: Second Semester

Course Introduction and Description

Introduction

This course – Linear Algebra II –is set to teach us the study of linear maps on finite-dimensional vector spaces. Eventually we will learn what all these terms mean. We will define vector spaces and discuss their elementary properties before gradual introduction to finite-dimensional vector spaces.

Description

The status of this course is two units. It is designed in two Modules i.e. Module 1 and Module 2 respectively. Module 1 is grouped into two study sessions, and module 2 is grouped into two study sessions as well. Therefore, the course can be summarised as having 2 modules and 4 study sessions in all.

This Course Guide gives a brief summary of the contents of the course material: vector spaces (or linear spaces), finite dimensional vector spaces, linear transformations, characteristic polynomials and characteristic equation are will be our focus.

i. COURSE PREREQUISITES

You should note that although this course has no subject pre-requisite, you are expected to have:

1. Satisfactory level of English proficiency
2. Basic Computer Operations proficiency

3. Online interaction proficiency
4. Web 2.0 and Social media interactive skills
5. MATH 102

ii. COURSE LEARNING RESOURCES

i. Course Textbooks

You will be provided with the following materials; Course Guide Study Modules. In addition, the course comes with a list of recommended textbooks, which though are not compulsory for you to acquire, but are necessary as supplements to the course material.

iii. COURSE OBJECTIVES AND OUTCOME

After studying this course, you should be able to:

1.Identify and solve related problem in finite dimensional vector space(linear span, linearly dependent and linearly independent, basis and dimension).

2. Identify and solve related problem in linear transformation and characteristic polynomial and characteristic equation.

3.Define special matrices: diagonal, triangular, and symmetric prove important result in each module.

iv. ACTIVITIES TO MEET COURSE OBJECTIVES

Specifically, this course shall comprise of the following activities:

1. Studying courseware
2. Listening to course audios
3. Watching relevant course videos
4. Field activities, industrial attachment or internship, laboratory or studio work (whichever is applicable)
5. Course assignments (individual and group)
6. Forum discussion participation
7. Tutorials (optional)
8. Semester examinations (CBT and essay based).

v. TIME (TO COMPLETE SYLABUS/COURSE)

This course requires three hours weekly and span over a period of eleven (11) weeks

vi. GRADING CRITERIA AND SCALE

Grading Criteria

A. Formative assessment

Grades will be based on the following:

Individual assignments/test (CA 1,2 etc)	20
Group assignments (GCA 1, 2 etc)	10
Discussions/Quizzes/Out of class engagements etc	10

B. Summative assessment (Semester examination)

CBT based	30
Essay based	30
TOTAL	**100%**

C. Grading Scale

A = 70-100
B = 60 – 69
C = 50 - 59
D = 45-49
F = 0-44

D. Feedback

Courseware based:

1. In-text questions and answers (answers preceding references)
2. Self-assessment questions and answers (answers preceding references)

Tutor based:

1. Discussion Forum tutor input
2. Graded Continuous assessments

Student based:

1. Online programme assessment (administration, learning resource, deployment, and assessment).

vii. LINKS TO OPEN EDUCATION RESOURCES

OSS Watch provides tips for selecting open source, or for procuring free or open software.

SchoolForge and SourceForge are good places to find, create, and publish open software. SourceForge, for one, has millions of downloads each day.

Open Source Education Foundation and Open Source Initiative, and other organisation like these, help disseminate knowledge.

Creative Commons has a number of open projects from Khan Academy to Curriki where teachers and parents can find educational materials for children or learn about Creative Commons licenses. Also, they recently launched the School of Open that offers courses on the meaning, application, and impact of "openness."

Numerous open or open educational resource databases and search engines exist. Some examples include:

- OEDb: over 10,000 free courses from universities as well as reviews of colleges and rankings of college degree programmes
- Open Tapestry: over 100,000 open licensed online learning resources for an academic and general audience
- OER Commons: over 40,000 open educational resources from elementary school through to higher education; many of the elementary, middle, and high school resources are aligned to the Common Core State Standards
- Open Content: a blog, definition, and game of open source as well as a friendly search engine for open educational resources from MIT, Stanford, and other universities with subject and description listings

- Academic Earth: over 1,500 video lectures from MIT, Stanford, Berkeley, Harvard, Princeton, and Yale
- JISC: Joint Information Systems Committee works on behalf of UK higher education and is involved in many open resources and open projects including digitising British newspapers from 1620-1900!

Other sources for open education resources

Universities

- The University of Cambridge's guide on Open Educational Resources for Teacher Education (ORBIT)
- OpenLearn from Open University in the UK

Global

- Unesco's searchable open database is a portal to worldwide courses and research initiatives
- African Virtual University (http://oer.avu.org/) has numerous modules on subjects in English, French, and Portuguese
- https://code.google.com/p/course-builder/ is Google's open source software that is designed to let anyone create online education courses
- Global Voices (http://globalvoicesonline.org/) is an international community of bloggers who report on blogs and citizen media from around the world, including on open source and open educational resources

Individuals (which include OERs)

- Librarian Chick: everything from books to quizzes and videos here, includes directories on open source and open educational resources
- K-12 Tech Tools: OERs, from art to special education
- Web 2.0: Cool Tools for Schools: audio and video tools
- Web 2.0 Guru: animation and various collections of free open source software
- Livebinders: search, create, or organise digital information binders by age, grade, or subject (why re-invent the wheel?)

viii. ABU DLC ACADEMIC CALENDAR/PLANNER

	PERIOD											
Semester	Semester 1				Semester 2				Semester 3			
Activity	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEPT	OCT	NOV	DEC
Registration												
Resumption												
Late Registn.												
Facilitation												
Revision/ Consolidation												
Semester Examination												

N.B: - All Sessions commence in January

- 1 Week break between Semesters and 6 Weeks vocation at end of session.

- Semester 3 is OPTIONAL (Fast-tracking, making up carry-overs & deferments)

ix. COURSE STRUCTURE AND OUTLINE

Course Structure

WEEK	MODULE	STUDY SESSION	ACTIVITY
Week1& 2	**STUDY MODULE 1**	**Study Session 1** Title: Vector spaces (or linear spaces) Pp. 17	1. Read Courseware for the corresponding Study Session 2. View the Video(s) on this Study Session 3. Listen to the Audio on this Study Session 4. View any other Video/U-tube (address/sitehttps://bit.ly/1Oyhc6M) 5. View referred Animation (Address/Sitehttps://bit.ly/1Oyhc6M)
Week3 & 4		**Study Session 2** Title: Finite dimensional vector spaces Pp. 29	1. Read Courseware for the corresponding Study Session 2. View the Video(s) on this Study Session 3. Listen to the Audio on this Study Session 4. View any other Video/U-tube (address/sitehttps://bit.ly/2CEKVOo) 5. View referred Animation (Address/Sitehttps://bit.ly/2CEKVOo)
Week 5& 6	**STUDY**	**Study Session 1** Title: Linear Transformations Pp. 42	1. Read Courseware for the corresponding Study Session 2. View the Video(s) on this Study Session 3. Listen to the Audio on this Study Session 4. View any other Video/U-tube (address/site bit.ly/2OqaTtG) 5. View referred Animation (Address/Site bit.ly/2CCkdpK)
Week7 & 8		**Study Session 2** Title: Characteristic polynomials and	1. Read Courseware for the corresponding Study Session 2. View the Video(s) on this Study Session 3. Listen to the Audio on this Study Session

	MODULE 2	Characteristic equations. Pp. 75	4. View any other Video/U-tube (address/site bit.ly/2TZgXPP) 5. View referred Animation (Address/Site bit.ly/2Ue6Tlo)
Week 13		**REVISION/TUTORIALS (On Campus or Online)& CONSOLIDATION WEEK**	
Week 14& 15		**SEMESTER EXAMINATION**	

Course Outline

MODULE 1
Study Session 1: Vector spaces (or linear spaces)
Study Session 2: Finite dimensional vector spaces

MODULE 2
Study Session 1: Linear Transformations
Study Session 2: Characteristic Polynomials and Characteristic Equations

Study Modules

MODULE 1

Contents

Study Session 1

Vector Spaces (Or Linear Spaces)

Section and Subsection Headings

Introduction

I welcome you to study session one of this course. We begin this course by working through a vector space. A vector space, or linear space, is an algebraic structure that frequently provide a home for solutions of mathematical models while *Linear algebra* is a study of linear spaces and linear transformations between them. We call elements of a linear space, vectors.

It is vital for us to note that we can scale vectors and add vectors. When you scale two or more vectors and then add together, the resulting *linear combination* is another vector. In this study session, we introduce the concept of linear space (or vector space) linear subspaces and we consider some examples.

1.0 Study Session Learning Outcome

After studying this session, I expect you to be able to:

1. Explain the concept of linear space, linear subspaces
2. Distinguish between the linear space and the linear subspaces with examples
3. Distinguish between Sums and Direct Sums
4. Prove the Direct Sums propositions

2.0 Main Content

2.1. Linear Spaces

Definition: A ***linear space*** (or ***vector space***) is a set V together with field of *scalars* F and two functions (operations):addition $+ : V\times V\rightarrow V$ and (scalar) multiplication $\cdot : F \times V\rightarrow V$ such that

(a) $x + y = y + x$ for x, y in V
(b) $x + (y + z) = (x + y) + z$ for $x,y,z \in V$
(c) there is an element $0 \in V$ such that for each $x \in V$, $x + 0 = x$
(d) for each $x \in V$ there corresponds $y \in V$ such that $x + y = 0$
(e) for each $\alpha,\beta \in F$ and $x \in V$, $\alpha(\beta x) = (\alpha\beta)x$
(f) for each $\alpha,\beta \in F$ and $x \in V$, $(\alpha + \beta)x = \alpha x + \beta x$
(g) for each $\alpha \in F$ and $x,y \in V$, $\alpha(x + y) = \alpha x + \alpha y$, and
(h) $1x = x$ for each $x \in V$

Remarks

i. If $x + y = x + z$ then $y = z$. This ***cancellation*** law holds because $y = 0 + y = (-x + x) + y = -x + (x + y) = -x + (x + z) = (-x + x) + z = 0 + z = z$.

ii. $0x+x = (0+1)x= x$ so $0x = 0$ (we are using the same symbol 0 for the zero element of F and V, since it tends not to cause confusion).

iii. The zero element 0 of (b) and y, the additive inverse of x, in (c) are unique; we usually write $-x$ for the additive inverse of x. Since $-1x + x = -1x + 1x = (-1 + 1)x= 0x = 0$ then $-x = -1x$.

iv. For us, the scalar field F will always be either the real numbers R or complex numbers C. However, there are other fields, for example, the rational numbers Q. There are finite fields, e.g., Zp(the integer's mod p) where p is a prime number. Vector spaces over finite fields find applications in, for instance, cryptography but we will not consider them in this course.

Examples of linear spaces:

(a) R^n: a vector $x \in R^n$ is an n-tuple of real numbers, $x = (x_1,x_2...x_n)$. The set of real numbers R is the scalar field. Vector addition and scalar multiplication are defined component-wise: if x and y are vectors and $\alpha \in R$ then $x + y = (x_1,x_2,...,x_n) + (y_1,y_2,...,y_n) := (x_1 + y_1,x_2 + y_2,...,x_n+ y_n)$ and $\alpha x = \alpha(x_1,x_2,...,x_n) := (\alpha x_1,\alpha x_2,...,\alpha x_n)$.

(b) C^n: a vector $z \in C^n$is an n-tuple of complex numbers, $z = (z_1,z_2...z_n)$. The set of complex numbers C is the scalar field. Vector addition and scalar multiplication are defined component-wise: if w and z are vectors and $\alpha \in C$ then $w + z = (w_1,w_2,...,w_n) + (z_1,z_2,...,z_n) := (w_1 + z_1,w_2 + z_2,...,w_n+ z_n)$ and $\alpha z = \alpha(z_1,z_2,...,z_n) := (\alpha z_1,\alpha z_2,...,\alpha z_n)$.

(c) $M_{m\times n}$: a vector is a rectangular $m \times n$ array of real numbers. Vector addition and scalar multiplication are defined entry-wise: If A and B are in $M_{m\times n}$ and $\alpha \in R$ then $A + B$ and αA are defined to be the $m \times n$ matrices whose entry in the i^{th} row and j^{th} column are $(A + B)(i,j) := A(i,j) + B(i,j)$ and $(\alpha A)(i,j) := \alpha A(i,j)$.

For example, $M_{2\times 2}$:

$$M_{2\times 2} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{R} \right\}$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a+e & b+f \\ c+g & d+h \end{pmatrix} \text{and} \ \alpha \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \alpha a & \alpha b \\ \alpha c & \alpha d \end{pmatrix}$$

Note: allowing entries and scalars to be complex numbers turns $M_{m\times n}$ into a complex vector space.

(d) $C(I;\mathrm{R})$, I an interval of real numbers: a vector is any continuous real valued function $f : I \to \mathrm{R}$. The scalars are real numbers. Functions f and g are added and scaled "component-wise" or pointwise, i.e., $f + g$ and αf are functions whose value at $x \in I$ is $(f + g)(x) := f(x) + g(x)$ and $(\alpha f)(x) := \alpha f(x)$.

(e) Π: a vector is any polynomial on R. Vector addition and scalar multiplication are defined pointwise, as in $C(I;\mathrm{R})$.

(f) Π_n: a vector is any polynomial of degree $n$on R. Vector addition and scalar multiplication are defined pointwise, as in $C(I;\mathrm{R})$.

(g) $BC(\mathrm{R};\mathrm{R})$: a vector is any bounded continuous function f: R → R. Addition and scalar multiplication are defined pointwise.

(h) l^2: a vector x is a sequence of real numbers that is square sum-able: so $x \in l^2$ if and only if $x = (x_1,x_2...,x_k...)$ and $^{\mathrm{X}}x^2_i < \infty$. Addition and scalar multiplication are as usual, defined component-wise.$i\geq 1$

(i) l^p(for $p \geq 1$): a vector x is a sequence $x = (x_1,x_2,...,x_k,...)$ of real numbers such that and $^{\mathrm{X}}|x_i|^p<\infty.i\geq 1$ Addition and scalar multiplication are, as usual, defined component-wise.

(j) l^∞: a vector x is a bounded sequence of real numbers. More precisely, $x = (x_1,x_2,...,x_k,...) \in l^\infty$ if and only if there is an $M > 0$ (a bound) such that $|x_i| <$

M for each $i = 1,2,\ldots.$ Addition and scalar multiplication are again, defined component-wise.

(k) **c**: a vector x is a convergent sequence of real numbers. Addition and scalar multiplication are defined component-wise.

(l) $\mathbf{c}_0$: a vector x is a convergent sequence of real numbers with limit 0. Addition and scalar multiplication are defined component-wise.

Remark: By replacing R with C, the above examples become examples of *complex* vector spaces.

We often move around in a vector space by scaling and adding vectors from a *special* collection of vectors. For example, in R^2 we can generate the vector (2,2) by the combination 2(1,0) + 2(0,1). In fact, we can generate any vector in R^2 from the two vectors (1,0) and (0,): $(a,b) = a(1,0)+b(0,1)$. By contrast, the two vectors (1,3) and (−2,−6) generate only vectors along the line described by the equation $y = 3x$.

2.1.1 Subspaces of Linear Spaces

You will realise that sometimes, subsets of a linear space Vis itself vector space, when using the addition and scalar multiplication inherited from the parent space V. For example, the set $S \subset R^2$ of all point on the horizontal axis is a vector space under the usual vector addition and scalar multiplication of R^2

Definition: a subset S of a linear space Vis a ***subspace*** of V, provided S is a vector space using vector addition and scalar multiplication as they are defined in V.

Examples of subspaces:

(a) any plane in R^3 containing the origin is a subspace of R^3.

(b) $\mathbf{c}_0$is a subspace of **c**.

(c) $BC(R;R)$ is a subspace of $C(R;R)$.

(d) Π_n is a subspace of Π, which in turn, is a subspace of $C(R;R)$

(e) define the set $D := \{f \in C([a,b];\mathrm{R}) \mid f^0$ exists and is continuous on $[a,b]\}$. D is a subspace of $C([a,b];\mathrm{R})$

(f) the set 0} consisting of just the zero vector of V, is always a subspace of V

(g) if S and T are two subspaces of L then so is S ∩ T, the set of all vectors in V common to *both* S and T

Theorem 1. A subset **S** of a linear space V is a subspace of V if and only if $\alpha x + \beta y \in$ **S** whenever x, y ∈**S** and α, β are scalars.

In-text Question 1:

Is W a subspace of V (where W consists of all matrices with zero determinant and V is a vector space of all 2×2 matrices over real field)?

In-text answer 1:

1. *W is not a subspace of V, this is because if we let* $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ *and* $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. *Then, A and B belongs to W. but* $A + B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ *does not belong to W. Since det(A+B)*$\neq 0$

2.2 Sums and Direct Sums

We will find that the notions of vector space sums and direct sums are useful. We define these concepts here. Suppose $U_1,\ldots,U_m$ are subspaces of V. The ***sum*** of $U_1,\ldots,U_m$,

denoted $U_1 + \cdots + U_m$, is defined to be the set of all possible sums of elements of $U_1,\ldots,U_m$. More precisely, $U_1 + \cdots + Um = \{ u_1 + \cdots + u_m : u_1 \in U_1, \ldots, u_m \in U_\mathrm{m}\}$.

You should verify that if $U_1,\ldots,U_m$ are subspaces of V, then the sum $U_1 + \cdots + U_m$ is a subspace of V.

Let us consider some examples of sums of subspaces. Suppose U is the set of all elements of F^3 whose second and third coordinates equal 0, and W is the set of all elements of F^3, whose first and third coordinates equal 0:

$U = \{(x,0,0) \in F^3 : x \in F\}$ and $W = \{(0,y,0) \in F^3 : y \in F\}$.

Then,

$U + W = \{(x,y,0) : x,y \in F\}$. 1.1

As another example, suppose U is as above and W is the set of all elements of F^3 whose first and second coordinates equal each other, and whose third coordinate equals 0:

$W = \{(y, y, 0) \in F^3 : y \in F\}$.

Then $U + W$ is also given by1.1.

Let us suppose that $U_1,...,U_m$ are subspaces of V. Clearly $U_1,...,U_m$ are all contained in $U_1 + \cdots + U_m$ (to see this, consider sums $u_1 + \cdots + u_m$ where all except one of the u's are 0). Conversely, any subspace of V containing $U_1,...,U_m$ must contain $U_1 + \cdots + U_m$ (as sub spaces must contain all finite sums of their elements). Thus $U_1 + \cdots + U_m$ is the smallest subspace of V containing $U_1,...,U_m$.

Again, suppose $U_1,...,U_m$ are subspaces of V, such that $V = U_1 + \cdots + U_m$. This entails that, every element of V could be written in the form

$u_1 + \cdots + u_m$, where each $u_j \in U_j$.

We will be especially interested in cases where each vector in V can be uniquely represented in the form above. This situation is so important that we give it a special name: direct sum. Specifically, we say that V is the *direct sum* of subspaces $U_1,...,U_m$, written $V = U_1 \oplus \cdots \oplus U_m$, if each element of V can be written uniquely as a sum $u_1 + \cdots + u_m$, where each $u_j \in U_j$.

Let us again consider some examples of direct sums. Suppose U is the subspace of F^3 consisting of those vectors whose last coordinate equals 0 and W is the subspace of F^3 consisting of those vectors whose first two coordinates equal 0:

$U = \{(x,y,0) \in F^3 : x,y \in F\}$ and $W = \{(0,0,z) \in F^3 : z \in F\}$.

Then $F^3 = U \oplus W$, you should verify.

Another example, suppose U_j is the subspace of F^n consisting of those vectors whose coordinates are all 0, except possibly in the j^{th} slot (for example, $U_2 = \{(0,x,0,...,0) \in F^n: x \in F\}$). Then,

$F^n = U1 \oplus \cdots \oplus Un$, you should verify.

A final example, consider the vector space P(F) of all polynomials with coefficients in F. Let U_e denote the subspace of P(F) consisting of all polynomials p of the form $p(z) = a_0 + a_2z^2 + \cdots + a_{2m}z^{2m}$,

and let U_o denote the subspace of P(F) consisting of all polynomials p of the form $p(z) = a1z + a3z3 + \cdots + a2m+1z2m+1$.

Here, m is a nonnegative integer and $a_0,...,a_{2m}+1 \in F$ (the notations U_e and U_o should remind you of even and odd powers of z). You should verify that.

$P(F) = U_e \oplus U_o$.

Sometimes none examples add to our understanding as much as examples. Consider the following three subspaces of F^3:

$U_1 = \{(x,y,0) \in F^3: x,y \in F\}$;

$U_2 = \{(0,0,z) \in F^3: z \in F\}$;

$U_3 = \{(0,y,y) \in F^3: y \in F\}$.

Clearly $F^3 = U_1 + U_2 + U_3$ because an arbitrary vector $(x,y,z) \in F^3$ can be written as

(x,y,z) = (x,y,0) + (0,0,z) + (0,0,0).

Where the first vector on the right side is in U_1, the second vector is in U_2, and the third vector is in U_3. However, F^3 does not equal the direct sum of U_1,U_2,U_3because the vector (0,0,0) can be written in two different ways as a sum $u_1+u_2+u_3$, with each $u_j \in U_j$. Specifically, we have (0,0,0) = (0,1,0) + (0,0,1) + (0,−1,−1) and, of course,

(0,0,0) = (0,0,0) + (0,0,0) + (0,0,0).

Where the first vector on the right side of each equation above is in U_1, the second vector is in U_2, and the third vector is in U_3.

In the example above, we show that something is not a direct sum by presenting that 0 does not have a unique representation as a sum of appropriate vectors. The definition of direct sum requires that every vector in the space have a unique representation as an appropriate sum. Suppose we have a collection of subspaces whose sum equals the whole space. The next proposition shows that when deciding whether this collection of subspaces is a direct sum, we need to only consider whether 0 can be uniquely written as an appropriate sum.

Proposition 2

Suppose that $U_1,...,U_n$ are subspaces of V. Then $V = U_1 \oplus \cdots \oplus U_n$ if both of the following conditions hold:

(a) $V = U1 + \cdots + Un$;

(b) the only way to write 0 as a sum $u_1 + \cdots + u_n$, where each $u_j \in U_j$, is by taking all the u_j's equal to 0.

Proof: First, suppose that $V = U_1 \oplus \cdots \oplus U_n$. Clearly, (a) holds (because of how sum and direct sum are defined). To prove (b), suppose that $u_1 \in U_1,...,u_n \in U_n$ and $0 = u1 + \cdots + un$.

Then, each u_j must be 0 (this follows from the uniqueness of the definition of direct sum because $0 = 0 + \cdots + 0$ and $0 \in U_1 ..., 0 \in U_n$), proving (b).

Now, suppose that (a) and (b) hold. Let $v \in V$. By (a), we can write $v = u_1 + \cdots + u_n$ for some $u_1 \in U_1$ ... $u_n \in U_n$. To show that this representation is unique, suppose that we also have $v = v_1 + \cdots + v_n$, where $v_1 \in U_1,...,v_n \in U_n$. Subtracting these two equations, we have $0 = (u_1 - v_1) + \cdots + (u_n - v_n)$.

Clearly, $u_1 - v_1 \in U_1,...,u_n - v_n \in U_n$, so the equation above and (b) imply that each $u_j - v_j = 0$. Thus, $u_1 = v_1,...,u_n = v_n$, as desired.

The next proposition gives a simple condition for testing which subspaces give a direct sum. You should note that this proposition deals only with the case of two subspaces. When asking about a possible direct sum with more than two subspaces, it is not enough to test that any two of the subspaces intersect only at 0. To see this, you have to consider the non-example presented just before 1.8.

In that non-example, we had $F^3 = U_1 + U_2 + U_3$, but F^3 did not equal the direct sum of U_1, U_2, U_3. However, in that non-example, we have $U_1 \cap U_2 = U_1 \cap U_3 = U_2 \cap U_3 = \{0\}$ (again, you should verify). The next proposition shows that with just two subspaces we get a necessary and sufficient condition for a direct sum.

Proposition 3: Suppose that U and W are subspaces of V. Then,
$V = U \oplus W$ if $V = U + W$ and $U \cap W = \{0\}$.

Proof: First, suppose that $V = U \oplus W$. Then $V = U + W$ (by the definition of direct sum). Also, if $v \in U \cap W$, then $0 = v + (-v)$, where.
$v \in U$ and $-v \in W$. By the unique representation of 0 as the sum of a vector in U and a vector in W, we must have $v = 0$. Thus, $U \cap W = \{0\}$, completing the proof in one direction.

To prove the other direction, now suppose that $V = U + W$ and $U \cap W = \{0\}$. To prove that $V = U \oplus W$, suppose that $0 = u + w$, where $u \in U$ and $w \in W$. To complete the proof, we need only show that $u = w = 0$ (by 1.8). The equation above implies that $u = -w \in W$. Thus, $u \in U \cap W$, and hence, $u = 0$. This, along with the equation above, implies that $w = 0$, completing the proof.

In-text Question 2:
If $V = U \oplus W$ is the sum unique?

In-text Answer 2:
The sum $V = U \oplus W$ is unique. I. e, it can be written in one way.

3.0 Tutor Marked Assignments (Individual or Group)

1) Suppose u and v belong to a vector space V. Simplify each of the following expressions:
   (a) $E_1 = 3(2u - 4v) + 5u + 7v$, (c) $E_3 = 2uv + 3(2u + 4v)$
   (b) $E_2 = 3u - 6(3u - 5v) + 7u$, (d) $E_4 = 5u - \frac{3}{v} + 5u$
2) Let $V = R^3$. Show that W is not a subspace of V, where;
   (a) $W = \{(a, b, c): a \geq 0\}$, (b) $W = \{(a, b, c): a^2 + b^2 + c^2 \leq 1\}$.

4.0 Conclusion/Summary

In this study session, we have covered the following:

1. The concept of linear space and linear subspaces with some examples.
2. The concept of Sums and Direct Sums with some examples.
3. We Prove the Direct Sums propositions.

5.0 Self -Assessment Questions

1) Let V be the vector space of functions $f: R \to R$. Which of the following subset is a subspace of V.
   (a) $W_1 = \{f(x): f(1) = 0\}$, all functions whose value at 1 is 0.
   (b) $W_2 = \{f(x): f(3) = f(1)\}$, all functions assigning the same value to 3 and 1.
   (c) $W_3 = \{f(t): f(-x) = -f(x)\}$, all odd functions.
2) Let U and W be subspaces of a vector space V. State any three properties of the direct sun $U + W$.

Answer to Self-Assessment Questions:

1. (a) W_1 is a subspace of V
   (b) W_2 is a subspace of V
   (c) W_3 is a subspace of V
2. (a) $U + W$ is a subspace of V
   (b) U and W are contained in $U + W$

(c) $U + W$ is the smallest subspace containing U and W; that is, $U + W =$ span (U, W)

6.0 Additional Activities (Videos, Animations & Out of Class Activities)

a. Watch the videos: http://wwwkhanacademy.org/math/linear-algebra/vectors_and_spaces/subspace_basis/v/linear-agebra-basis-of-a-subspaces?utm_source=YT&utm_medium=Desc&utm-campaign=LinearAlgebra

7.0 References/Further Reading

Axler, S. et al (1997). Linear Algebra Done Right. 2nd Edition. New York: Springer-verlag

“MATH212, Linear Algebra IINational Open University of Nigeria, (2009) at www.nou.edu,ng

Lipschutz, S.(2009).Linear Algebra.(Schaum’s Outline Series) 4th Edition. 2 Pennsylvania Plaza New York City. McGraw-Hill

Study Session 2

Finite Dimensional Vector Spaces

Section and Subsection Headings

Introduction

You are welcome to study session two. In study session 1, we introduced you to the concept of linear spaces. We discussed the vector spaces and subspaces and the sums and direct sums.

In this study session, we shall proceed with the finite dimensional vector spaces.

1.0 Learning Outcome

After studying this session, I expect you to be able to:

1. Explain the concept of linear span, linearly dependent and linearly independent
2. Distinguish between the linearly dependent and the linearly independent with examples
3. Describe the concept of basis and dimension
4. Prove the theorem related to above concepts

2.0 Main Content

2.1 Linear Spans

Let us launch this way! Suppose $u_1, u_2, ..., u_m$ are any vectors in a vector space V. Any vector of the form $a_1u_1 + a_2u_2 + \ldots + a_m u_m$, where the a_i are scalars, is called a ***linear combination*** of $u_1, u_2, ..., u_m$.

The collection of all such linear combinations, denoted by

Span $(u_{1,} u_{2,} ...,u_m)$ or span$(u_i)$is called the ***linear span*** of $u_1, u_2, ..., u_m$.

Clearly, the zero vector 0 belongs to span (u_i), because$0 =0u_1 +0u_2 + \ldots+0u_m$

Furthermore, suppose v and v^1belong to span(u_i), say, $v= a_1u_1 + a_2u_2 + \ldots + a_m u_m$ and $v^1= b_1u_1 + b_2u_2 + \ldots +b_m u_m$. Then, $v + v^1 = (a_1+ b_1) u_1 + (a_2 + b_2) u_2 + \ldots + (a_m + b_m)u_m$ and, for any scalar $k \in K$, $kv= ka_1u_1 + ka_2u_2 + \ldots +ka_mu_m$

Thus, $v + v^1$and kv also belong to span(u_i). Accordingly, span(u_i) is a subspace of V.

Generally, for any subset S of V, span(S) consists of all linear combinations of vectors in S or, when $S = \emptyset$, span(S)={0}. Thus, in particular, S is a spanning set of span(S).

Following theorem, which was partially proved above, holds.

Theorem 4: Let S be a subset of a vector space V. Then,

(i) Linear span L(S) is a subspace of V that contains S.

(ii) If W is a subspace of V containing S, then linear span L(S) W.

Condition (ii) in the above theorem may be interpreted as saying that linear span L(S) is the ‘‘smallest’’ subspace of V containing S.

Proof

i. We show that L(S) $\neq \emptyset$ and isclosed under vector addition and scalar multiplication.

Notice that if $v \in S$, then 1.v= $v \in L(S) \Longrightarrow S \subset L(S)$ and L(S) $\neq \emptyset$.

Suppose that $u, v \in L(S)$. Then,

$$\left.\begin{aligned} u &= \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \\ v &= \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n \end{aligned}\right\} \alpha_i, \beta_i \in K$$

Then$u + v = (\alpha_1 + \beta_1)v_1 + (\alpha_2 + \beta_2)v_2 + \cdots + (\alpha_n + \beta_n)v_n \in L(S)$ because $u + v$ is a linear combination ofv_i.

Also,$\alpha u = \alpha(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = \alpha\alpha_1 v_1 + \alpha\alpha_2 v_2 + \cdots + \alpha\alpha_n v_n \Longrightarrow$ $L(S)$ is a subspace of V.

ii. Suppose W is a subspace of V containing S, then all multiples $\alpha_1 v_1, \alpha_1 v_1, \ldots, \alpha_1 v_1$ $\in$ W and so, the sum

$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ $\in$ W

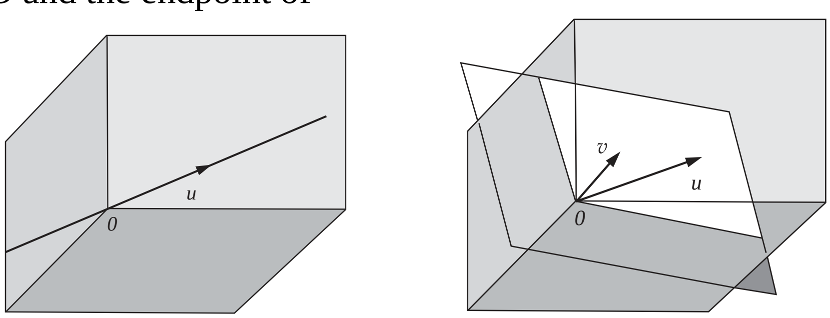
This means that W contains all linear combination of vectors in S.

$\Longrightarrow L(S) \subset$ W

Since W is arbitrary, this implies that $L(S)$ is the smallest subspace of V containing S.

Example 1 Consider the vector space V = R^3.

(a) Let u be any nonzero vector in R^3. Then span(u)consists of all scalar multiples of u, as shown in Fig. 1.2.1: (a). u. Geometrically, span(u)is the line through the origin O and the endpoint of



(a) (b)

Figure 1.2.1: Vector Spaces

(b) Let u and v be vectors in R^3 that are not multiples of each other. Then span(u, v) is the plane through the origin O and the endpoints of u and v as shown in Fig. 1.2.1(b).

(c) Consider the vectors $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 =(0, 0, 1)$in R^3. Recall Example 1.2.1(a) that every vector in R^3 is a linear combination of e_1, e_2, e_3. That is, e_1, e_2, e_3 form a spanning set of R^3. Accordingly, span $(e_1,e_2,e_3) \in R^3$.

2.2 Linear Dependence and Independence

Let V be a vector space over a field K. You will note that the following defines the notion of linear dependence and independence of vectors over K. (One usually suppresses mentioning K when you understand the field.) This concept plays an essential role in the theory of linear algebra and in mathematics in general.

Definition: we say that the vectors $v_1, v_2, ..., v_m$ in V are ***linearly dependent*** if there exist scalars $a_1, a_2, ..., a_m$ in K, not all of them 0, such that$a_1v_1 + a_2v_2 + \dots + a_mv_m = 0$

Otherwise, we say that the vectors are ***linearly independent***.

We may restate the above definition as follows. Consider the vector equation$x_1v_1 + x_2v_2 + \dots + x_mv_m = 0$ (*)

Where the x's are unknown scalars. This equational ways has the zero solution $x_1= 0, x_2= 0, \dots,x_m= 0$, Suppose this is the only solution, that is, suppose we can show $x_1v_1 + x_2v_2 + \dots +x_mv_m = 0$ implies $x_1 =0,x_2 =0, ..., x_m=0$.

Then the vectors $v_1, v_2, ...,v_m$ are linearly independent. On the other hand, suppose the equation (*) has a nonzero solution; then the vectors are linearly dependent.

A set $S =\{v_1, v_2, ...,v_m\}$of vectors in V is ***linearly dependent*** or ***independent*** according to whether the vectors $v_1, v_2, ..., v_m$ are linearly dependent or independent.

An infinite set S of vectors is ***linearly dependent*** or ***independent*** according to whether there exist or not, vectors $v_1, v_{2,} ...,v_k$ in S that are linearly dependent.

The following remarks follow directly from the above definition.

1. Suppose 0 is one of the vectors v_1, v_2, ...,v_m , say v_1= 0. Then the vectors must be linearly dependent, because we have the following linear combination where the coefficient of $v_1 \neq 0$:$1v_1 + 0v_2 + \ldots + 0v_m = 1.0 + 0 + \ldots + 0 = 0$
2. Suppose v is a nonzero vector. Then v, by itself, is linearly independent, because Kv = 0; v $\neq$0 implies k =0

3. Suppose two of the vectors v_1, v_2, ...,v_m are equal or one is a scalar multiple of the other, say $v_1 = kv_2$. Then the vectors must be linearly dependent, because we have the following linear combination where the coefficient of $v_1 \neq 0$:v_1 - kv_2+ 0v3+ … + $0v_m$ = 0

4. Two vectors v_1 and v_2 are linearly dependent if one of them is a multiple of the other.

5. If the set {$v_{1,}$ ...,v_m}is linearly independent, then any rearrangement of the vectors $\{v_{i_1}, v_{i_2}, \ldots, v_{i_m}\}$ is also linearly independent.

6. If a set S of vectors is linearly independent, then any subset of S is linearly independent. Alternatively, if S contains a linearly dependent subset, then S is linearly dependent.

Example2

1. Let u = (1, 1, 0), v = (1, 3, 2), w = (4, 9, 5). Then u, v, w are linearly dependent, because

   3u +5v - 2w =3(1, 1, 0) +5 (1, 3, 2) -2(4, 9, 5)= (0, 0, 0) =0
2. We show that the vectors $u = (1,2,3)$, $v = (2,5,7)$, $w = (1.3,5)$ are linearly independent. We form the vector equation $xu + yv + zw = 0$, where x,y,z are unknown scalars. This yields

$$x\begin{bmatrix}1\\2\\3\end{bmatrix}+y\begin{bmatrix}2\\5\\7\end{bmatrix}+z\begin{bmatrix}1\\3\\5\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}\text{or}\begin{bmatrix}x+2y+z\quad=0\\2x+5y+3z=0\\3x+7y+5z=0\end{bmatrix}\text{or}\begin{bmatrix}x+2y+z=0\\x+z=0\\2z=0\end{bmatrix}$$

Back-substitution yields x =0, y =0, z =0. We have shown that

xu+ yv + zw =0 implies x =0, y =0, z =0

Accordingly, u, v, w are linearly independent.

In-text Question 1

When is a vector said to be linearly independent and / or linearly dependent?

In-text answer 1

S is said to be linearly dependent if there exists scalars $c_1, c_2, \ldots, c_n$*not all zero such that* $c_1u_1 + c_2u_2 + \cdots + c_nu_n = 0$. *Otherwise, it is linearly independent.*

2.3 Basis and Dimension

First, we state two equivalent ways to define a basis of a vector space V.

Definition A: A set S ={u_1, u_2, ...,u_n} of vectors is a ***basis*** of V if it has the following two properties: (1) S is linearly independent. (2) S spans V.

Definition B: A set S ={u_1, u_2, ..., u_n} of vectors is a ***basis*** of V if every v∈ V. can be written uniquely, as a linear combination of the basis vectors.

The following is a fundamental result in linear algebra.

Theorem 5: Let V be a vector space such that one basis has m elements and another basis has n elements. Then $m = n$.

Definition: The ***dimension*** of a vector space V is the number of non-zero vector in the basis. We may say that a vector space V has finite dimension n or n-dimensional, written dim$V = n$.

If V has a basis with n elements. Theorem 4.12 tells us that all bases of V have the same number of elements, so this definition is well defined. The vector space {0} is defined to have dimension 0.Suppose a vector space V does not have a finite basis. Then V is said to be of infinite dimension or to be infinite-dimensional.

The above fundamental Theorem is a consequence of the following ‘‘replacement lemma’’

Lemma 1: Suppose $\{v_1,v_2,...,v_n\}$ spans V, and suppose $\{w_1,w_2,...,w_m\}$ is linearly independent. Then $m \leq n$, and V is spanned by a set of the form$\{w_1,w_2,...,w_m, v_{i1};v_{i2},...,v_{in-m}\}$. Thus, in particular, n +1 or more vectors in V are linearly dependent.

Observe in the above lemma that we have replaced m of the vectors in the spanning set of V, by the m independent vectors and still retained a spanning set.

Examples3

This subsection presents important examples of bases of some of the main vector spaces appearing in this study manual.

(a) Vector space K^n: consider the following n vectors in K^n:$e_1 =(1,0,0,0,...,0,0)$, $e_2 =(0,1,0,0,...,0,0)$, ..., $e_n =(0,0,0,0,...,0,1)$

These vectors are linearly independent. (For example, they form a matrix in echelon form.) Furthermore, any vector $u =(a_1,a_2,...,a_n)$ in K^n can be written as a linear combination of the above vectors. Specifically,$v =a_1 e_1 + a_2 e_2+ ... + a_ne_n$

Accordingly, the vectors form a basis of K^n which we call *the **usual** or **standard** **basis** of K^n*. Thus (as one might expect), K^n has dimension n. In particular, any other basis of K^n has n elements.

(b)Vector space $M =M_{r;s}$ of all $r \times s$ matrices: The following six matrices form a basis of the vector space $M_{2;3}$ of all 2×3 matrices over K:

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

Generally, in the vector space $M =M_{r;s}$ of all $r \times s$ matrices, let E_{ij} be the matrix with ij-entry 1 and 0’s elsewhere. Then all such matrices form a basis of $M_{r;s}$ called the usual or standard basis of $M_{r;s}$. Consequently, dim $M_{r;s}$=rs.

(C) Vector space $P_n(t)$ of all polynomials of degree $\leq$ n: The set $S =\{1,t,t^2,t^3,...,t^n\}$ of n +1 polynomials is a basis of P_n(tÞ). Specifically, we express any polynomial f

(t) of degree n as a linear combination of these powers of t, and one can show that these polynomials are linearly independent. Therefore, dim $P_n(t) = n + 1$.

The following two algorithms, which we essentially describe here, find such a basis (and hence, the dimension) of W.

Algorithm 4.1 (Row space algorithm)

Step 1.Form the matrix M whose rows are the given vectors.

Step 2.Row reduce M to echelon form.

Step 3.Output the nonzero rows of the echelon matrix.

Sometimes we want to find a basis that only comes from the original given vectors. The next algorithm accomplishes this task.

Algorithm 4.2 (Casting-out algorithm)

Step 1.Form the matrix M whose columns are the given vectors.

Step 2.Row reduce M to echelon form.

Step 3.For each column C_k in the echelon matrix without a pivot, delete (cast out) the vector u_k from the list S of given vectors.

Step 4. Output the remaining vectors in S (which correspond to columns with pivots).

We emphasise that in the first algorithm, we form a matrix whose rows are the given vectors, whereas in the second algorithm we form a matrix whose columns are the given vectors.

Example 4 Let W be the subspace of R^5 spanned by the following vectors:

$u_1 = (1,2,1,3,2)$, $u_2 = (1,3,3,5,3)$, $u_3 = (3,8,7,13,8)$,

$u_4 = (1, 4, 6, 9, 7)$, $u_5 = (5, 13, 13, 25, 19)$

Find a basis of W consisting of the original given vectors, and find dim W.

Solution

Form the matrix M whose columns are the given vectors, and reduce M to echelon form:

$$M = \begin{bmatrix} 1 & 1 & 3 & 1 & 5 \\ 2 & 3 & 8 & 4 & 13 \\ 1 & 3 & 7 & 6 & 13 \\ 3 & 5 & 13 & 9 & 25 \\ 2 & 3 & 8 & 7 & 19 \end{bmatrix} \text{This reduces to} \begin{bmatrix} 1 & 1 & 3 & 1 & 5 \\ 0 & 1 & 2 & 2 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The pivots in the echelon matrix appear in columns C_1, C_2, C_4. Hence, we ‘‘cast out’’ the vectors u_3 and u_5 from the original five vectors. The remaining vectors u_1, u_2, u_4, which correspond to the columns in the echelon matrix with pivots, form a basis of W. Thus, in particular, dim W = 3.

Remark: In the above example, we justify the casting-out algorithm, but we repeat it again here for emphasis. The fact that column C_3 in the echelon matrix in Example 4.13 does not have a pivot means that the vector equation-$xu_1 + yu_2 = u_3$has a solution hence,u_3 is a linear combination of u_1 and u_2.

Similarly, the fact that C_5 does not have a pivot means that u_5 is a linear combination of the preceding vectors. We have deleted each vector in the original spanning set that is a linear combination of preceding vectors. Thus, the remaining vectors are linearly independent and form a basis of W.

Example 5V is a subspace R^4generated by {(1,1,-1,2), (2,0,-1,3), (1,-1,0,1)}. Find a homogeneous system whose solution space is V.

Solution: if v=(x,y,z,t) ∈ V, then v = a +b + c where a,b,c are scalars.

Therefore;

$$a + 2b + c = x$$
$$a \quad - c \quad = y$$

$$-a - b \quad = z$$
$$2a + 3b + c = t$$

Reducing the augmented matrix to echelon yields

$$\begin{pmatrix} 1 & 2 & 1:x \\ 1 & 0 & -1:y \\ -1 & -1 & 0:z \\ 2 & 3 & 1:t \end{pmatrix} \text{Reduces to} \begin{pmatrix} 1 & 2 & 1:x \\ 0 & -2 & -1:y-x \\ 0 & 1 & 1:z+x \\ 0 & -1 & -1:t-2x \end{pmatrix}$$

$$\text{Reduces to} \begin{pmatrix} 1 & 2 & 1:x \\ 0 & -2 & -2:y-x \\ 0 & 0 & 0:z+y+2z \\ 0 & 0 & 0:3x+y-2t \end{pmatrix}$$

Therefore, the required homogeneous system is

$$\begin{aligned} z+y+2z &= 0 \\ 3x+y-2t &= 0 \end{aligned}$$

In-text Question 2

When do we say a set S isa basis for V and what is a dimension of V?

In-text Answer 2

S is a basis for V if S spans V and S is linearly independent. The dimension of V of a vector space V is defined as the number of nonzero vectors in a basis for V.

3.0 Tutor Marked Assignments (Individual or Group)

1. Find the dimension and a basis of the solution space W of each homogeneous system:

(a) x + y + 2z = 0
2x + 3y + 3z = 0
x + 3y + 5z = 0

(b) x + 2y + z - 2t = 0
-2x + 4y + 4z - 3t = 0
3x + 6y + 7z - 4t = 0

(c) x + 2y + 2z - s + 3t = 0
x + 2y + 3z + s + t = 0
3x + 6y + 8z + s + 5t = 0

2. Find a homogeneous system whose solution set W is spanned by $\{u_1, u_2, u_3\} = \{(1, 2, 0, 3), (1, 1, 1, 4), (1, 0, 2, 5)\}$.

4.0 Conclusion/Summary

In this study session, we have covered the following:

1. The concept of linear span, linearly dependent and linear independent.
2. Distinguished between the linear dependent and the linear independent with examples.
3. Concept of basis and dimension with examples.
4. Prove some of the theorem related to linearly dependent and linear independent, basis and dimension.

5.0 Self-Assessment Questions

1. Determine whether or not each of the following lists of vectors in R^3 is linearly dependent:

(a) $u_1 = (1, 2, 5)$, $u_2 = (1,3,1)$, $u_3 = (2, 5, 7)$, $u_4 = (3,1,4)$,

(b) $u = (1, 2,5)$, $v = (2, 5, 1)$, $w = (1, 5, 2)$,

(c) $u = (1, 2, 3)$, $v = (0, 0, 0)$, $w = (1, 5, 6)$.

2. Determine whether or not each of the following form a basis of R^3:

(a) $(1, 1, 1), (1, 0, 1)$;

(b) $(1, 2, 3),\ (1, 3, 5),\ (1, 0, 1),\ (2, 3, 0)$;

(c) $(1, 1, 1), (1, 2, 3), (2, 1, 1)$;

(d) $(1, 1, 2), (1, 2, 5), (5, 3, 4)$.

Answer to Self-Assessment Questions:

1. (a) The given vectors are linearly dependent since
$$-39u_1 - 26u_2 + 31u_3 + u_4 = 0.$$
(b) The given vectors are linearly independent since
$$xu + yv + zw = 0 \implies x = y = z = 0.$$
(c) The given vectors are linearly dependent since for any real number a
$$0u + av + 0w = 0.$$

2. (a) The vectors $(1, 1, 1), (1, 0, 1)$ do not form basis for R^3
   (b) The vectors (1, 2, 3), (1 ,3, 5), (1, 0, 1), (2, 3, 0) do not form basis for R^3
   (c) The vectors $(1, 1, 1), (1, 2, 3), (2, 1, 1)$ form basis for R^3
   (d) The vectors $(1, 1, 2), (1, 2, 5), (5, 3, 4)$ do not form basis for R^3

6.0 Additional Activities (Videos, Animations &Out of Class Activities) e.g.

a. Watch the Videos: http://wwwkhanacademy.org/math/linear-algebra/vectors_and_spaces/linear_independence/v/span-and-linear-indepence-example?utm_source=YT&utm_medium=Desc&utm-campaign=LinearAlgebra

7.0 References/Further Reading

Axler, S. et al (1997). Linear Algebra Done Right. 2nd Edition. New York: Springer-verlag

“MTH212, Linear Algebra II, National Open University of Nigeria, 2009 at www.nou.edu,ng

Lipschutz, S. (2009).Linear Algebra.4th Edition (Schaum’s Outline Series). 2 Pennsylvania Plaza New York City:McGraw-Hill

MODULE 2

Content

Study Session 1

Linear Transformations

Section and Subsection Headings

3.0 Tutor Marked Assignments (Individual or Group Assignments)
4.0 Study Session Summary and Conclusion
5.0 Self-Assessment Questions
6.0 Additional Activities (Videos, Animations & Out of Class Activities)
7.0 References/Further Reading

Introduction

You have already learnt about a vector space and several concepts related to it. In this study session, we initiate the study of certain mappings between two vector spaces, which we call linear transformations. You can say that the importance of these mappings is realised from the fact that, in the calculus of several variables, every continuously differentiable function is replaceable, to a first approximation, by a linear one. This fact is a reflection of a general principle that every problem on the change of some quantities under the action of several factors is regarded to a first approximation, as a linear problem. It often turns out that this gives an adequate result. Also in physics, it is important for us to know how vectors behave under a change of the coordinate system. This requires a study of linear transformations.

In this study session, we study linear transformations and their properties, as well as two spaces associated with linear transformations and their properties, and two spaces associated with their dimensions. Then, we prove the existence of linear transformations with some specific properties; we will also discuss the notion of an isomorphism between two vector spaces, which allows us to say that all finite dimensional vector spaces of the same dimension are the "Same", in a certain sense.

Finally, we state and prove the Fundamental Theorem of Homomorphism and some of its corollaries, and apply them to various situations.

1.0 Study Session Learning Outcome

After studying this session, I expect you to be able to:

1. Verify the linearity of certain mappings between vector spaces
2. Construct linear transformations with certain specified properties
3. Calculate the rank and nullity of a linear operator
4. Define an isomorphism between two vector spaces
5. Show that two vector spaces are isomorphic if they have the same dimension
6. Prove and use the Fundamental Theorem of Homomorphism
7. Make matrix representation of a linear map

2.0 Main Content

2.1 Linear Transformations

By now, you are familiar with vector spaces IR^2 and IR^3. Here, we shall consider the mapping f: R^2□R^3 defined by f(x,y) = (x,y,0) (see fig. 1 below).

f is a well-defined function. In addition, you will notice that

(i) f((a, b) + (c, d)) = f((a + c, b + d)) = (a + c, B + d,0) = (a, b, 0) + (c,d,0)

f(a, b) +f(c ,d), for (a,b), (c,d)∈ R^2 and Z

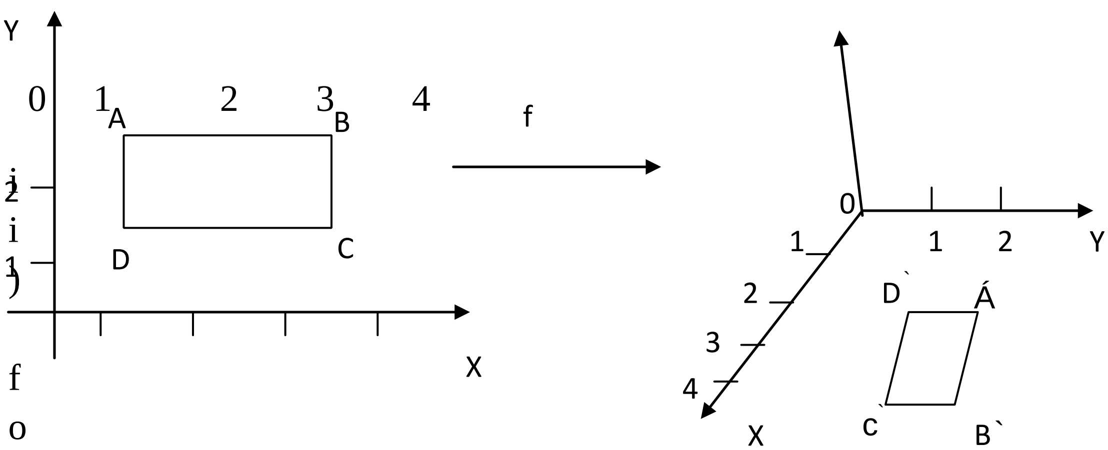


Fig. 2.11: f transforms ABCD to A`B`C`D`.

(ii) for any α € R and (a,b) ∈ R^2, f((αa, αb)) = (αa, αb, 0) = α (a,b,0) = αf((a,b)).

Therefore, we have a function f between two vector spaces such that (i) and (ii) above hold true.

(i) Says that the sum of two plane vectors is mapped under f to the sum of their images under f.

(ii) Says that a line in the plane $R^2$is mapped under f to a line in R^2.

The properties (i) and (ii) together say that f is linear, a term that we now define.

Definition: Let U and V be vector spaces over a field F. A linear ***transformation*** (or ***linear operator***) from U to V is a function T: U V, such that

LT1:$T(u_1 + u_2) = T(u_1) + T(u_1)$, for U_1, $U_2 \in U$, and

LT2:$T(\alpha u) = \alpha T(u)$ for $\alpha \in F$ and $U \in U$.

We can combine the conditions LT1 and LT2 to give the following equivalent condition. LT3:$T(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 T(u_1) + \alpha_2 T(u_2)$, for $\alpha_1, \alpha_2 \in F$ and $u_1, u_2 \in U$.

What we are saying is that [LT1 and LT2] ⇔LT3. We can easily show this as follows:

We will show that LT3 ⟹LT1 and LT3 ⟹ LT2. Now, LT3 is true $\forall$ $\alpha_1, \alpha_2 \in F$.

Therefore, it is certainly true for α_1, α_2 ,that is, LT1 holds.

Again, to show that LT2 is true, consider $T(\alpha u)$ for any $\alpha \in F$ and $u \in U$. We have $T(\alpha u) = T(\alpha u + 0.\ u) = \alpha T(u) + 0.T(u) = \alpha T(u)$, thus proving that LT2 holds.

Before going further, let us note two properties of any linear transformation T:U0 → V, which follow from LT1 (or LT2, or LT3).

LT4:T(0) = 0. Let us see why this is true. Since T(0) = T(0 + 0) = T(0) +T(0) (by LT1, we subtract T(0) from both sides to get T(0) = 0.

LT5:T(-u) = - T(u) ~~V~~ u ∈ U. why is this so? Well, since 0 = T(0) = T(u – u) = T(u) + T(-u), we get T(-u) = - T(u).

Now let us look at some common linear transformations.

Example 1: Consider the vector space U over a field F, and the function T: U → U defined by T(u) = u for all u ∈ U.

Show that T is a linear transformation. (This transformation is called the **identity transformation,** and is denoted by I_u, or Just I, if the underlying vector space is understood).

Solution: For any α, β ∈ F and u_1, u_2∈ U, we have

$T(\alpha u_1 + \beta u_2) = \alpha u_1 + \beta u_2 = \alpha T(u_1) + \beta T(u_2)$

Hence, LT3 holds, and T is a linear transformation.

Example 2: Let T: U → V be defined by T(u) = 0 for all u ∈ U.

Check that T is a linear transformation. (we refer to it as the null or **Zero Transformation,** and is denoted by **0**).

Solution: For any α,β,∈ F and u_1, u_2∈ U, we have $T(\alpha u_1 + \beta u_2) = 0 = \alpha.0 + \beta.0 = \alpha T(u_1) + \beta T(u_2)$.

Therefore, T is linear transformation.

Example 3: Consider the function $pr_1:R^n \rightarrow R$, defined by $pr_1[(x_1, \ldots\ldots\ldots x_n)] = x_1$. Shows that this is a linear transformation. (This is called the projection on the first coordinate. Similarly, we can define $pr_1: R^n \rightarrow R$ by $pr_1 [(x_1 \ldots X_i\text{-}1, X_i, \ldots\ldots\ldots, x_n)] = x_1$ to be the **projection** on the i^{th} **Coordinate** for i = 2, ...,n. For instance, $pr_2:R^3 \rightarrow R$:

$pr_2(x,y,z) = y$.)

Solution: We will use LT3 to show that projection is a linear operator. For $\alpha, \beta, \in$ F and $(x_1 \ldots\ldots\ldots x_n)$, $(y_1 \ldots\ldots\ldots\ldots y_n)$ in R^n, we have

$Pr_1 [\alpha(x_1 \ldots\ldots\ldots\ldots x_n) + \beta(y_1 \ldots\ldots\ldots\ldots y_n)]$

$= pr_1 (\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2 \ldots\ldots\ldots\ldots \alpha x_n + \beta y_n) = \alpha x_1 + \beta y_1 = \alpha pr_1, [(x_1, \ldots\ldots, x_n)] + \beta pr_1 [(y_1, \ldots\ldots, y_n)].$

Thus, pr1 (and similarly pr1) is a linear transformation.

Before going to the next example, we make a remark about projections.

Remark: consider the function $p: R^3 \rightarrow R^2$: $p(x,y,z) = (x,y)$. This is a projection from R^3 on to the xy-plane. Similarly, the functions f and g, from $R^3 \rightarrow R^2$, defined by $f(x,y,z) = (x,z)$ and $g(x,y,z) = (y,z)$ are projections from R^3 onto the xz-plane and the yz-plane, respectively.

In general, any function: $R^n \rightarrow R^m$ (n> m), which is defined by dropping any (n – m) coordinates, is a projection map.

Now let us see another example of a linear transformation that is very geometric in nature.

Example 4: Let T: $R^2: \rightarrow R^2$ be defined by $T(x,y) = (x, -y) \ \forall \ x,y \in R$. Show that T is a linear transformation. (we show this **reflection** in the x-axis in fig 2).

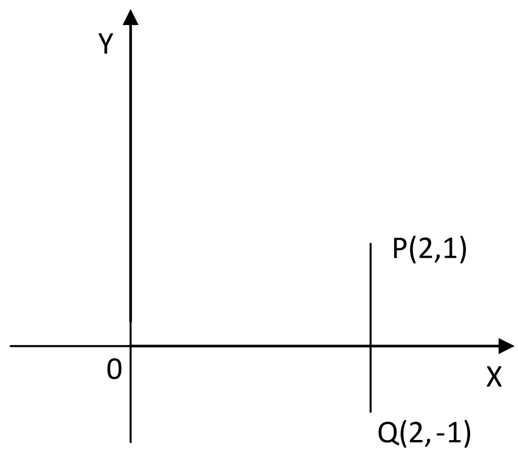


Fig 2.1.2: Q is the reflection of P in the X-axis.

Solution: For $\alpha, \beta \in F$ and $(x_1, y_1), (x_2, y_2) \in R^2$, we have
$T[\alpha (x_1,y_1) + \beta (x_2,y_2) = T (\alpha x_1 + \beta x_2, \alpha y_1 + \beta y_2) = (\alpha x_1 + \beta x_2, - ay_1 - \beta y_2)$
$= \alpha (x1, - y1) + \beta (x2, - y2)$
$= \alpha T (x_1, y_1) + \beta T(x_2, y_2)$.
Therefore, T is a linear transformation.

So far we have given examples of linear transformations. Now we give an example of a very important function that is not linear. This example's importance lies in its geometric applications.

Example 5: Let u_0 be a fixed non-zero vector in U. Define $T: U \to U$ by $T(u) = u + u_0 \forall u \in U$. Show that T is not a linear transformation. (We call T the translation by u_0. See Fig 2.1.3 for a geometrical view).

Solution: T is not a linear transformation since LT4 does not hold. This is because $T(0) = u_0 \neq 0$

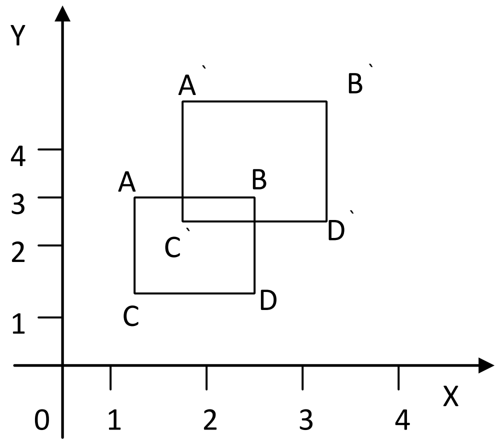


*Fig. 2.1.3: A`B`C`D` is the translation of ABCD by (1,1)***.**

Example 6: Let W be a subspace of a vector space U over a field F. W gives rise to the quotient space U/W. Consider the map $T:U\rightarrow U/W$ defined by $T(u) = u + W$. Show that T is a linear transformation.

Solution: for $\alpha, \beta \in$ **F** and $u_1, u_2 \in U$ we have $T(\alpha u_1 + \beta u_2) = (\alpha u_1 + \beta u_2) + W = (\alpha u_1 + W) + (\beta u_2 + W)$. $= \alpha (u_1 + W) + \beta (u_2 + W)$

$= \alpha T(u_1) + \beta T(u_2)$

Thus, T is a linear transformation.

You have already seen that a linear transformation $T:U \rightarrow V$ must satisfy $T(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 T(u_1) + \alpha_2 T(u_2)$, for $\alpha_1, \alpha_2 \in F$ and $u_1, u_2 \in U$. Generally, we can show that,

LT6: $\mathbf{T(\alpha_1 u_1 + \ldots + \alpha_n u_n) = \alpha_1 T(u_1) + \ldots. + \alpha_n T(u_n)}$, where $\alpha_1 \in F$ and u_1, U.

Let us show this by induction, that is, we assume the above relation for n = m, and prove it for m + 1.Now, $T(\alpha u_1 + \ldots + \alpha_m u_m + \alpha_{m+1} u_{m+1})$

$= T(u + \alpha_{m+1} u_{m+1})$, where $u = \alpha_1 u_1 + \ldots + \alpha_m u_m$

$= T(u) + \alpha_{m+1} T(u_{m+1})$, since the result holds for n = 2

$= T(\alpha_1 u_1 + \ldots + \alpha_m u_m) + \alpha_{m+1} T(u_{m+1})$

$\alpha_1 T(u_1) + \ldots + \alpha_m T(u_m) + \alpha_{m+1} T(u_{m+1})$, since we have assumed the result for n = m.

Thus, the result is true for n = m+1. Hence, by induction, it holds true for all n.

Let us now come to a very important property of any linear transformation $T:U\rightarrow V$. In study session 2 of module we mentioned that every vector space has a basis. Thus, U has a basis. We will now show that T is completely determined by its values on a basis of U. More precisely, we have -

Theorem 1: Let S and T be two linear transformation from U to V, where $\dim_1 U = n$. Let $(e_1 \ldots e_n)$ be a basis of U. Suppose $S(e_1)$ i = 1, …, n. Then S(u) = T(u) for all $u \in U$.

Proof: Let $u \in U$. Since $(e_1, \ldots, e_n)$ is a basis of U, u can be uniquely written as $u = \alpha_1 e_1 + \ldots + \alpha_{\mathbf{n}} e_{\mathbf{n}}$, where the α_1 are scalars.

Then, $S(u) = S(\alpha_1 e_1 + \ldots + \alpha_{\mathbf{n}} e_n)$

$= \quad \alpha_1 S(e_1) + \ldots + \alpha_n S(e_n)$, by LT6

$= \quad \alpha_1 T(e_1) + \ldots + \alpha_n T(e_n)$

$= \quad \alpha_1(a_1 e_1 + \ldots + \alpha_n e_n)$, by LT6

$= \quad T(u)$.

What we have just proved is that once we know the values of T on a basis of U, then we can find T(u) for any $u \in U$.

Note: Theorem 1 is true even when U is not finite – dimensional. The proof, in this case, is on the same line as above.

Let us see how the idea of Theorem 1 helps us to prove the following useful result.

Theorem 2: Let V be a real vector space and **T: R**$\rightarrow$ V be a linear transformation. Then there exists $v \in V$ such that $T(\alpha) = \alpha v \; \cancel{\forall} \alpha \in \mathbf{R}$.

Proof: A basis for R is (1). Let $T(1) = V \in V$. Then, for any $\alpha \in R$, $T(\alpha) = \alpha T(1) = \alpha v$.

Once you have read you will realise that this theorem says that T(R)is a vector space of dimension one, whose basis is [T(1)].

Theorem 3: Let (e_1 …., e_n) be a basis of U and let v_1, …,v_n be any n vectors in V. Then there exists one linear transformation T:U→ V such that $T(e_i) = v_i$, i = 1, …, n.

Proof: Let u ∈ U. Then u can be uniquely written as $u = \alpha_1 e_1 + \ldots, \alpha_n e_n$.
Define $T(u) = \alpha_1 v_1 + \ldots + \alpha_n v_n$. The T defines a mapping from U to V such that $T(e_1) = v_1 \cancel{\forall} i = 1, \ldots, n$. Let us now show that T is linear. Let a, b be scalars and u, u\` ∈ U. The ∃ scalar $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$ such that $u = \alpha_1 e_1 + \ldots + \alpha_n e_n$ and u\` = $\beta_1 e_1 + \ldots + \beta_n e_n$.
Then, au + bu\` = $(a\alpha_1 + b\,\beta_1)\, e_1 + \ldots + (a\,\alpha_{\mathbf{n}} + b\,\beta_{\mathbf{n}})\, e_{\mathbf{n}}$.
Hence, T(au + bu\`) = $(a\alpha_1 + b\,\beta_1)\, v_1 + \ldots + (a\alpha_n + b\,\beta_{\mathbf{n}})\, v_{\mathbf{n}}$
$= a(\alpha_1 v_1 + \ldots + \alpha_{\mathbf{n}} v_{\mathbf{n}}) + b(\beta_1 v_1 + \ldots + \beta_{\mathbf{n}} v_{\mathbf{n}})$
= aT(u) + bT(u\`)

Therefore, T is a linear transformation with the property that $T(e_i) = v_i \cancel{\forall}$ i. Theorem 1 now implies that T is the only linear transformation with the above properties.
Let us see how Theorem 3 can be used.

Example 7: $e_1 = (1,0,0)$. $e_2 = (0, 1, 0)$ and $e_3 = (0, 0, 1)$ form the standard basis of R^3.
Let (1,2), (2,3) and (3,4) be three vectors in R^2. Obtain the linear transformation T: $R^3 \to R^2$ such that T(e1) = (1,2), $T(e_2) = (2,3)$ and $T(e_3) = (3,4)$.

Solution: By Theorem 3 we know that ∃ $T:R^3 \to R^2$ such that $T(e_1) = (1,2)$, $T(e_2) = (2,3)$, and $T(e_3) = (3,4)$. We want to know what T(x) is, for any

$x = (x_1, x_2, x_3) \in R^3$, Now, $X = x_1 e_1 + x_2 e_2 + x_3 e_3$.

Hence, $T(x) = x_1 T(e_1) + x_2 T(e_2) + x_3 T(e_3)$
$= x_1 (1,2) + x_2 (2,3) + x_3 (3,4)$
$= (x_1 + 2x_2 + 3x_3, 2x_1 + 3x_2 + 4x_3)$

Therefore, $T(x_1, x_2, x_3) = (x_1 + 2x_2 + 3x_3, 2x_1 + 3x_2 + 4x_3)$ is the definition of the linear transformation T.

2.1.1 Spaces Associated with a Linear Transformation

In study session 1 of module 1, you find that given any function, there is a set associated with it, namely, its range. We will now consider two sets that are associated with any linear transformation, T. These are the range and the kernel of T.

2.1.2 The Range Space and the Kernel

Let U and V be vector spaces over a field F. Let $T:U\rightarrow V$ be a linear transformation. We will define the range of T as well as the Kernel of T. At first, you will see them as sets. We will prove that these sets are also vector spaces over F.

Definition: The ***range*** of T, denoted by R(T), is the set $\{T(x)/ x \in U\}$. The **kernel** (or null space) of **T**, denoted by **Ker T,** is the set $\{x \in U/ T(x) = 0\}$ Note that $R(T) \subseteq V$ and $\text{Ker } T \subseteq U$.

To clarify these concepts, consider the following examples.

Example 8: Let $I: V \rightarrow V$ be the identity transformation (see Example 1). Find R(I) and Ker I.

Solution: $R(I) = \{I(v) \mid v \in V\} = \{v | v \in V\} = V$. Also, $\text{Ker } I = \{v \in V | I(v) = 0\} = \{v \in V | \ v = 0\} = \{0\}$.

Example 9: Let $T: R^3\rightarrow R$ be defined by $T(x_1, x_2, x_3) = 3x_1 + x_2 + 2x_3$FindR(T) and Ker T.

Solution: $R(T) = \{x \in R | \exists\ x_1, x_2, x_3, \in R \text{ with } 3x_1 + x_2 + 2x_3. = x\}$. For example, $0 \in R(T)$. Since $0 = 3.0 + 0 + 2.0 = T(0, 0, 0,)$

Also, $I \in R\ (T)$, since $I = 3.1/3 + 0 + 2.0 = T(1/3, 0, 0)$, or $I = 3.0 + 1 + 2.0 = T(0, 1, 0)$, or $I = T(0, 0, ½)$ or $I = T\ (1/6, ½, 0)$.

Now can you see that R(T) is the whole real line R? This is because, for any $\alpha \in$ R, $\alpha = \alpha\ 1 = \alpha\ T(1/3,0,0) = T(^{\alpha}/_{3}, 0, 0) \in R\ (T)$.

$\text{Ker } T = \{(x_1, x_2, x_3) \in R^3 |\ 3x_1 + x_2 + 2x_3 = 0\}$.

For example, $(0,0,0) \in$ Ker T. But $(1, 0,0) \notin$ Ker T . :. Ker T $\neq R^3$. In fact, Ker T is the plane $3x_1 + x_2 + 2x_3 = 0$ in R^3.

Example 10: let $T:R^3 \rightarrow R^3$ be defined by $T(x_1,x_2, x_3) = (x_1 - x_2 + 2x_3, 2x_1 + x_2, -x_1 - 2x_2 + 2x_3)$. Find R(T) and Ker T.

Solution: to find R(T), we must find conditions on $y_1, y_2, y_3 \in R$ so that $(y_1, y_2, y_3) \in R(T)$, i.e. ., we must find some $(x_1, x_2, x_3) \in R^3$ so that $(y_1, y_2, y_3) = T(x_1, x_2, x_3)$

$=(x_1 - x_2 + 2x_3, 2x_1 + x_1 - 2x_2 + 2x_3)$.

This means $x_1 - x_2 + 2x_3=y_1$……………….. (1)

$2x + x_2 = y_2$ …………………………..(2)

$-x_1 = 2x_2 + 2x_3 = y_3$ …………………. (3)

Subtracting 2 times Equation (1) from Equation (2) and adding Equations (1) and (3) we have

$3x_2 - 4x_3 = y_2 + - 2y_1$ ………………… (4) and

$- 3x_2 + 4x_3 = y_1 + y_3$ ………………… (5)

Adding Equations (4) and (5), we have

$y_2 - 2y_1 + y_1 + y_3 = 0$, that is, $y_2 + y_3 = y_1$,

Thus, $(y_1, y_2, y_3 \in R\ (T) \Longrightarrow\ y_2 + y_3 = y_1$.

On the other hand, if $y2 + y3 = y1$. We can choose

$X_3 = 0,\ x_2 = \frac{y2 - 2y1}{3}$ and $x_1 = y_1 + \frac{y2 - 2y1}{3}\ = \frac{y1 + y2}{3}$

Then, we see that T $(x_1, x_2, x_3) = (y_1, y_2, y_3)$. Thus, $y_2 + y_3 = y_1 \Longrightarrow (y_1, y_2, y_3) \in R(T)$.

Hence, $R(T) = \{(y_1,y_2,y_3) \in R^3 / y_2 + y_3 = y_1\}$

Now $(x_1,x_2,x_3) \in$ Ker T if the following equations are true: $x_1 – x_2 +$
$2x_3 = 0$, $2x_1 + x_2 = 0$
$-x_2 – 2x_2 + 2x_3 = 0$

Of course $x_1 = 0$, $x_2 = 0$, $x_3 = 0$ is a solution. Are there other solutions? To answer this, we proceed as in the first part of this example. We see that $3x_2 = 4x_3 = 0$. Therefore, $x_3 = (3/4)\,x_2$.

Also, $2x_1 + x_2 = 0 \Longrightarrow x_1 = - x_2/2$.

Thus, we can give arbitrary values to x_2 and calculate x_1 and x_3 in terms of x_2. Therefore, Ker T = $\{(-\,^{\alpha}/_{2}, \alpha, (3/4)\alpha) : \alpha \in R\}$.

In this example, we see that finding R(T) and Ker T amounts to solving a system of equations. In study session2, you will learn a systematic way of solving a system of linear equations by the use of matrices and determinants.

Now that you are familiar with the sets R(T) and Ker T, we will prove that they are vector spaces.

Theorem 4: Let U and V be vector spaces over a field F. Let $T: U \rightarrow V$ be a linear transformation. Then Ker T is a subspace of U and R(T) is a subspace of V.

Proof: Let $x_1, x_2, \in$ Ker T $\subseteq$ U and $\alpha_1, \alpha_2 \in$ **F**. Now, by definition, $T(x_1) = T(x_2) = 0$. Therefore, $\alpha_1T(x_1) + \alpha_2T(x_2) = 0$

But $\alpha_1T(x_1) + \alpha_2T(x_2) = T(\alpha_1x_1 + \alpha_2 x_2)$.

Hence, T $(\alpha_1x_1 + \alpha_2x_2) = 0$. This means that
$\alpha_1x_1 + \alpha_2x_2 \in$ Ker T.

Thus, by Theorem 4 of study session 2, Ker T is a subspace of U.

Let $y_1, y_2 \in R(T) \subseteq V$, and $\alpha_1, \alpha_2 \in F$. Then, by definition of R(T), there exist $x_1, x_2 \in U$ such that $T(x_1) = y_1$ and $T(x_2) = y_2$.

So, $\alpha_1y_1 + \alpha_2y_2 = \alpha_1T(x_1) + \alpha_2T(x_2) = T(\alpha_1x_1 + \alpha_2x_2)$.

Therefore, $\alpha_1y_1 + \alpha_2y_2 \in R(T)$, which proves that R(T) is a subspace of V.

Now that we have proved that R(T) and Ker T are vector spaces, you know, from study session 2, that they must have a dimension. We will study these dimensions now.

2.1.4 Some types of Linear Transformations

Let us recall that there can be different types of functions, some of which are one-one, onto or invertible. We can also define such types of linear transformations as follows.

Definition: Let $T: U \to V$ be a linear transformation.

a) T is called ***one-one*** (or ***injective***) if, for $u_1, u_2 \in U$ with $u_1, \neq u_2$, we have $T(u_1) \neq T(u_2)$. If T is injective, we also say T is 1 – 1. Note that T is **1 – 1, if $T(u_1)$** $= T(u_2) \Rightarrow u_1 = u_2$.

b) T is called ***onto*** (or ***surjective***) if, for each $v \in V$, $\exists u \in U$ such that $T(u) = v$, that is $R(T) = V$.

Can you think of examples of such functions? The identity operator is both one-one and onto. Why is this so? Well, $I:V \to V$ is an operator, such that, if $v_1, v_2 \in V$ with $v_1 \neq v_2$ then $I(v_1) \neq I(v_2)$ Also, $R(I) = V$, so that I is onto. An important result that characterises injectivity is the following.

Theorem 5: $T:U \to V$ is one-one if Ker T = (0).

Proof: First assume T is one – one. Let $u \in$ Ker T. Then $T(u) = 0 = T(0)$. This means that $u = 0$. Thus, Ker T = (0). Conversely, let Ker T = (0). Suppose $u_1, u_2 \in U$ with $T(u_1) = T(u_2) \Rightarrow T(u_2 - u_2) = 0 \Rightarrow u_1 - u_2 \in \text{Ker } T \Rightarrow u_1 - u_2 = 0 \Rightarrow u_1 = u_2$. Therefore, T is 1– 1.

Suppose now that T is a one – one and onto linear transformation from a vector space U, to a vector space V. Then, (Theorem 4), we know that T^{-1}exist. Nevertheless, is T^{-1} linear? The answer to this question is 'yes', as is shown in the following theorem.

Theorem 6: Let U and V be vector space over a field F. Let $T:U \rightarrow V$ be a none-one and onto linear transformation. Then $T^{-1}: V \rightarrow U$ is a linear transformation. In fact, T^{-1} is also 1 – 1 and onto.

Proof: Let $y_1, y_2 \in V$ and $\alpha_1, \alpha_2 \in F$. Suppose $T^{-1}(y_1) = x_1$ and $T^{-1}(y_2) = x_2$. Then, by definition, $y_1 = T(x_1)$ and $y_2 = T(x_2)$.

Now, $\alpha_1 y_1 + \alpha_1 y_2 = \alpha_1 T(x_1) + \alpha_2 T(x_2) = T(\alpha_1 x_1 + \alpha_1 x_2)$

Hence, $T^{-1}(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 x_1 + \alpha_2 x_2 [T^{-1}(y) = x \Leftrightarrow T(x) = y]$

$= \alpha_1 T^{-1}(y_1) + \alpha_2 T^{-1}(y_2)$

This shows that T^{-1} is a linear transformation.

We will now show that T^{-1} is 1 -1, for this suppose, $y_1, y_2 \in V$ such that

$T^{-1}(y_1) = T^{-1}(y_2)$

Let $x_1 = T^{-1}(y_1)$ and $x_2 = T^{-1}(y_2)$.

Then $T(x_1) = y_1$ and $T(x_2) = y_2$. We know that $x_1 = x_2$. Therefore, $T(x_1) = T(x_2)$, that is, $y_1 = y_2$. Thus, we have shown that $T^{-1}(y_1) = T^{-1}(y_2) \implies y_1 = y_2$, proving that T^{-1} is 1 – 1. T^{-1} is also surjective because, for any $u \in U$, $\exists\ T(u) = v \in V$ such that $T^{-1}(v) = u$.

Theorem 8 says that a one-one and onto linear transformation is ***invertible***, and the inverse is a one-one and onto linear transformation.

This theorem immediately leads us to the following definition.

Definition. Let U and V be vector spaces over a field F, and let $T:U \rightarrow V$ be a one-one and onto linear transformation. We call the T an ***Isomorphism*** between U and V. In this case, we say that U and V are ***isomorphic vector spaces and*** we denote this by $U \approx V$.

An obvious example of isomorphism is the identity operator. Can you think of any other?

Using these properties of an isomorphism, we can get some useful results, like the following:

Theorem 7: Let $T: U \rightarrow V$ be an isomorphism. Suppose $\{e_1 \ldots e_n)$ is a basis of U. Then $\{T(e_1), \ldots, T(e_n)\}$ is a basis of V.

Proof: First we show that the set $\{T(e_1) \ldots, T(e_n)\}$ spans V. Since T is onto, R(T) = V. Thus, from E12 you know that $\{T(e_1),.., T(e_n)\}$ spans V.
Let us now show that $\{T(e_1),.., T(e_n)\}$ is linearly independent. Suppose there exist scalars $c_1,\ldots,c_n$, such that $c_1 T(e_1) + \ldots + c_n T(e_n) = 0$……………………………1
We must show that $c_1 = \ldots = c_n = 0$
Now, (1) implies that $T(c_1 e_1 + \ldots + c_n e_n) = 0$
Since T is one-one and T(0) = 0, we conclude that $c_1e_1 + \ldots + c_n e_n = 0$. But $\{e_1, \ldots,e_n\}$ is linearly independent. Therefore, $c_1 = \ldots = c_n = 0$.
Thus, we have shown that $\{T(e_1),\ldots, T(e_n)\}$ is a basis of V.
Remark: the argument showing the linear independence of $\{T(e_1), \ldots, T(e_n)\}$ in the above theorem, can be used to prove that any one-one linear transformation $T:U \to V$ maps any linearly independent subset of U, onto a linearly independent subset of V.
We now give an important result equating 'isomorphism' with '1-1' and with 'onto' in the finite-dimensional case.

Theorem 8: Let $T: U \to V$ be a linear transformation where U. V are of the same finite dimension. Then the following statements are equivalent.

a) T is 1 – 1
b) T is onto
c) T is an isomorphism

Proof: to prove the result, we will prove (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (a).
Let dim U = dim V = n.
Now (a) implies that Ker T = (0+ (from Theorem 7), Hence, nullity (T) = (0). Therefore, by Theorem 5, rank (T) = n that is dim R(T) = n = dim V. However, R (T) is a subspace of V. thus, by the remark following Theorem 12 of study session2, we get R (T) = V, i.e., T is onto, i.e., (b) is true. So (a) $\Rightarrow$ (b).

Similarly, if (b) holds then rank (T) = n, and hence, nullity (T) = 0 consequently, Ker T = {0}, and T is one-one. Where, T is one-one and onto, i.e.,t is an isomorphism. Therefore, (b) implies (c).

That (a) follows from (c) is immediate from the definition of an isomorphism.

With this, we have proven our result.

Caution: Theorem 10 is true for ***finite-dimensional spaces*** U and V, of the same ***dimension***. It is not true, otherwise. Consider the following counter-example.

Example 11: (To show that the spaces have to be finite-dimensional). Let V be the real vector space of all polynomials. Let D:V ⟶ V be defined by D ($a_0 + a_1$ x +...+ $a_r x^r$) = $a_1 + 2a_2 x + \ldots. + ra_r\, x^{r-1}$. Then, show that D is onto but not 1-1.

Solution: you should note that V has infinite dimension, a basis being {1, x, x^2…}. D is onto because any element of V is of the form $a_0 + a_1$ x + …. + $a_n x_n$ = D $a_0 x\left[\frac{a_1}{2}x^2 + \cdots + \frac{a_n}{n+1}x^2\right]$

D is not 1–1 because, for example, 1 ≠ 0 but D (1) = D (0) = 0.

If T: U ⟶ V is an isomorphism, then T maps a basis of U onto a basis of V. Therefore, dim U = dim V. In other words, if U and V are isomorphic then dim U = dim V. The natural question arises whether the converse is also true. That is, if dim U = dim V, both being finite, can we say that U and V are isomorphic? The following theorem shows that this is indeed the case.

Theorem 9: let U and V be finite-dimensional vector spaces over F. The U and V are isomorphic if dim U = dim V.

Proof: we have already seen that if U and V are isomorphic then dim U= dim V. Conversely, suppose dim U = dim V = n. We shall show that U and V are isomorphic. Let {e_1…, e_n} be a basis of U and {f_1….,f_n} be a basis of V. By Theorem 3, there exist a linear transformation T:U⟶ V such that T(e_1) = f_1, i = 1, ….., n. We shall show that T is 1-1.

Let $u = c_1e_1 + \ldots + c_n e_n$ be such that $T(u) = 0$ Then $0 =$
$T(u) = c_1 T(e_1) + \ldots + c_n T(e_n)$
$= c_1f_1 + \ldots + c_nf_n$.
Since $\{f_1,\ldots,f_n\}$ is a basis of V, we conclude that $c_1 = c_2 = \ldots = c_n = 0$. Hence, $u = 0$. Thus, Ker T = (0) and, by Theorem 7, T is one – one.
Therefore, by Theorem 10, T is an isomorphism, and U = V. An immediate consequence of this theorem follows.

Corollary 1: let V be a real (or complex) vector space of dimension n. Then V is isomorphic to R" (or C"), respectively.
Proof: since $\dim R_n = n = \dim_R V$, we get $V \approx R^n$. Similarly, if dim V = n, then $V \approx c_n$.

Remark: let V be a vector space over F and let $B = \{e_1\ldots, e_n\}$ be a basis of V. Each v € V can be uniquely expressed as $v = \sum_{i=1}^{n} \alpha_i e_i$. You must recall that $\alpha_1\ldots\alpha_n$ are coordinates of v with respect to B.
Define: $V \to F^n$: $\theta(v) = (\alpha_1 \ldots\ \alpha_n)$. Then θ is an isomorphism from V to F^n. This is because θ is 1-1, since the coordinates of v with respect to B are uniquely determined; thus, $V \approx F^n$.
Now, let us look at isomorphism between quotient spaces.

2.1.4 Homomorphism Theorems

We also call linear transformation ***vector space homomorphism***. There is a basic theorem that uses the properties of homomorphism to show the isomorphism of certain quotient spaces. It is simple to prove. However, it is very important because it is always in use to prove advanced theorems on vector spaces. (In the Abstract Algebra Course, we will prove this theorem in the setting of groups and rings).
Theorem 10: Let V and W be vector spaces over a field F and $T:V\to W$ be a linear transformation. Then $\mathbf{V/Ker\ T \approx R(T)}$.

Proof: you know that Ker T is a subspace of V, so that V/Ker T is a well-defined vector space over F. Also, R (T) = {T(v)\V ∈ V}. To proof the theorem, let us define θ V/Ker T ⟹ R (T) by θ (v + Ker T) = T (v).

First, we must show that θ is a well-defined function, that is, if v +Ker T = v` + Ker T then $\theta$ (v + Ker T) = $\theta$ (v` + Ker T),i.e. T(v) = T (v`).

Now, v + Ker T = v` + Ker T ⟹ (v –v`) ∈ Ker T (see study session 3, E23)

⟹T (v-v`) = 0 ⟹T(v) = T(v`) and hence, θ is well defined.

Next, we check that θ is a linear transformation. For this, let a, b∈ F and v. v` ∈V then,$\theta${a (v + Ker T) + b(v` + Ker T)}

= θ (av + bv` + Ker T) (ref. study session 3)

= T (av + bv`)

= aT(v) + bT (v`), since T is linear.

= a θ (v + Ker T) + b θ(v` + Ker T).Thus,

θ is a linear transformation.

We end the proof by showing that θ is an isomorphism. $\theta$is 1-1 (because θ (v + Ker T) = θ T(v) = 0 ⟹ v ∈ Ker T ⟹ v + Ker T = 0 (in V/Ker T).

Thus, Ker θ={0})

$\theta$is onto (because any element of R (T) is T (v) = θ (v) = θ (v + Ker T)). So we prove that θ is an isomorphism. This proves that V/Ker T = R (T).

Example 12: Show that $R^3/R \approx R^2$.

Solution: note that we can consider R as a subspace of R^3 for the following reason: any element a of R is equated with the element (α 0, 0) of R^3. Now, we define a function f: $R^3 \longrightarrow R^2$: f(α,β,γ) = (β,γ). Then, f is a linear transformation and Ker f = {(α, 0, 0) |α ∈ R} ≈ R. In addition, f is onto, since any element (α,β) of R^2 is f (0, α,β). Thus, by Theorem 13, $R^3/R \approx R^2$.

Note: in general, for any n ≥ m, $R^n/R^m \approx R^{n-m}$. Similarly, $C^{n-m} \approx C^n/C^m$ for n ≥ m.

Theorem 11. Let V be of finite dimension, and let F: V → U be linear. Then, dimV = dim(Ker F) + dim(ImF) = nullity(F) + rank(F)

Proof: since V is finite dimensional, let dimV = n and let dim(ker F) = K. if k = n. then V= ker F $\Longrightarrow$ F(v) = 0 $\forall$ v $\in$ V. Hence, imF = {0} and dim(imF) =0

Suppose that $1 \leq k \leq n$, we show thatdim(imF) $=$ n $-$ k

Let $\{v_1, v_2, \dots, v_k\}$ be a basis for ker F, we can extend this basis to that of V. i.e. S = $\{v_1, v_2, \dots, v_k, v_{k+1}, v_{k+2}, \dots, v_n\}$ is a basis for V.

We now prove that T = $\{F(v_{k+1}), F(vk +_2), \dots, F(v_n)\}$ is a basis for imF.

Let ω $\in$imF, then $\omega = F(v)$ for some v $\in$ V. Since S is a basis for V, we can write $v = c_1 v_1 + \ c_2 v_2 + \ \dots + c_k v_k + \ c_{k+1} v_{k+1} + \ c_{k+2} v_{k+2} + \ \dots + c_n v_n$

Where c_j are scalars.

$\omega = F(v)$

$= F(c_1 v_1 + c_2 v_2 + \ \dots + c_k v_k + c_{k+1} v_{k+1} + c_{k+2} v_{k+2} + \cdots + c_n v_n)$

Now

$$= c_1 F(v_1) + c_2 F(v_2) + \cdots + c_k F(v_k) + c_{k+1} F(v_{k+1}) + c_{k+2}(v_{k+2}) + \cdots + c_n F(v_{n)}$$

$= c_{k+1} F(v_{k+1}) + c_{k+2}(v_{k+2}) + \cdots + c_n F(v_{n)}$, since $F(v_{j)} = 0 \forall\, j = 1,2, \dots k$

Hence, T spans imF.

We finally show that T is linearly independent.

Suppose that

$$c_{k+1} F(v_{k+1}) + c_{k+2}(v_{k+2}) + \cdots + c_n F(v_{n)} = 0$$

Then,$F(c_{k+1} v_{k+1} + c_{k+2} v_{k+2} + \cdots + c_n v_n) = 0$

Hence, $c_{k+1} v_{k+1} + c_{k+2} v_{k+2} + \cdots + c_n v_n \in$ ker F and can be written as a linear combination of vectors in S.

That is $c_{k+1} v_{k+1} + c_{k+2} v_{k+2} + \cdots + c_n v_n = d_1 v_1 + d_2 v_2 + \ \dots + d_k v_k$

Where

That is $d_1 v_1 + d_2 v_2 + \ \dots + d_k v_k - c_{k+1} v_{k+1} - c_{k+2} v_{k+2} - \cdots - c_n v_n = 0$

Since S is independent, we see that

$d_1 = d_2 = \cdots = d_k = c_{k+1} = c_{k+2} = \cdots = c_n = 0$

So, T is linear independent and dim(imF) = 0

This complete the proof.

You would recall that we define the rank of a matrix A to be the dimension of its column space and row space. If we now view A as a linear mapping, then both definitions correspond, because the image of A is precisely its column space.

In-text Question 1
What are the basic condition(s) of linearity?

In-text Answer 1
The condition(s) for linearity are: $T(\alpha u) = AT(u)$ *and* $T(u_1 + u_2) = T(u_1) + T(u_2)$, *for all* $\alpha \in K$ *and for* $u_1, u_2 \in V$. *or equivalently,* $T(\alpha u_1 + \beta u_2) = \alpha T(u_1) + \beta T(u_2)$, $\alpha, \beta \in K$ *and for* $u_1, u_2 \in V$

2.2 Matrix Representation of a Linear Operator

Let T be a linear operator (transformation) from a vector space V into itself, and suppose $S = \{u_1, u_2, \ldots, u_n\}$ is a basis of V. Now, $T(u_1), T(u_2), \ldots, T(u_n)$ are vectors in V, and so each is a linear combination of the vectors in the basis S; say,

$$T(u_1) = a_{11} u_1 + a_{12}u_2 + \ldots + a_{1n}u_n$$
$$T(u_2) = a_{21} u_1 + a_{22}u_2 + \ldots + a_{2n}u_n$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$T(u_n) = a_{n1} u_1 + a_{n2}u_2 + \ldots + a_{nn}u_n$$

The following definition applies.

Definition: The transpose of the above matrix of coefficients, denoted by $m_S(T)$ or $[T]_S$, is called ***the matrix representation of T relative to the basis S***, or simply ***the matrix of T in the basis S***. (The subscript S may be omitted if the basis S is understood.)

Using the coordinate (column) vector notation, the matrix representation of T may be written in the form, $m_S(T) = [T]_S = [\,[T(u_1)]_S, [T(u_2)]_S, \ldots, [T(u_n)]_S]$
That is, the columns of m(T) are the coordinate vectors of $T(u_1), T(u_2), \ldots, T(u_n)$, respectively.

Example 13: Let $F:R^2 \longrightarrow R^2$ be the linear operator defined by $F(x, y) = (2x + 3y, 4x - 5y)$.

(a) Find the matrix representation of F relative to the basis $S = \{u_1; u_2\} = \{(1,2), (2,5)\}$.

(1) First find $F(u_1)$, and then write it as a linear combination of the basis vectors u_1 and u_2. (For notational convenience, we use column vectors.) We have

$$F(u_1)=F\left(\begin{bmatrix}1\\2\end{bmatrix}\right) = \begin{bmatrix}8\\-6\end{bmatrix} = x\begin{bmatrix}1\\2\end{bmatrix} + y\begin{bmatrix}2\\5\end{bmatrix} \text{and} \begin{matrix} x & + & 2y = & 8 \\ 2x & + & 5y = & -6 \end{matrix}$$

Solve the system to obtain $x = 52$, $y = -22$. Hence, $F(u_1) = 52u_1 - 22u_2$.

(2) Next find $F(u_2)$, and then write it as a linear combination of u_1 and u_2:

$$F(u_2)=F\left(\begin{bmatrix}2\\5\end{bmatrix}\right) = \begin{bmatrix}19\\-17\end{bmatrix} = x\begin{bmatrix}1\\2\end{bmatrix} + y\begin{bmatrix}2\\5\end{bmatrix} \text{and} \begin{matrix} x & + & 2y = & 19 \\ 2x & + & 5y = & -17 \end{matrix}$$

Solve the system to get $x = 129$, $y = -55$. Thus, $F(u_2) = 129u_1 - 55u_2$.

Now, write the coordinates of $F(u_1)$ and $F(u_2)$ as columns to obtain the matrix

$$[F]_S = \begin{bmatrix}52 & 129\\-22 & -55\end{bmatrix}$$

(b) Find the matrix representation of F relative to the (usual) basis $E = \{e_1, e_2\} = \{(1, 0), (0, 1)\}$.

Solution: find $F(e_1)$ and write it as a linear combination of the usual basis vectors, then write it as a linear combination of e_1 and e_2 and then find $F(e_2)$ and write it as a linear combination of the usual basis vectors, again, write it as a linear combination of e_1 and e_2. We have

$$F(e_1) = F(1, 0) = (2, 2) = 2e_1 + 2e_2$$

$$F(e_2) = F(0, 1) = (3, -5) = 3e_1 - 5e_2$$

and so $[F]_E = \begin{bmatrix}2 & 3\\2 & -5\end{bmatrix}$

You should note that the coordinates of $F(e_1)$and $F(e_2)$form the columns, not the rows of $[F]_E$. In addition, you would realise that the arithmetic is much simpler using the usual basis of R_2.

Example 14: let $F:R^2 \longrightarrow R^2$ be defined by F(x, y) = (2x +3y,4x -5y). Find the matrix representation $[F]_S$ of F relative to the basis S = {u1, u22} = {(1, 2), (2, 5)}.

Solution

(Step 1)First find the coordinates of (a, b) $\in R^2$ relative to the basis S. We have$\begin{bmatrix} a \\ b \end{bmatrix} = x\begin{bmatrix} 1 \\ -2 \end{bmatrix} + y\begin{bmatrix} 2 \\ -5 \end{bmatrix}$ or $\begin{matrix} x & + & 2y = & a \\ -2x & - & 5y = & b \end{matrix}$

Solving for x and y in terms of a and b yields x =5a +2b, y =2a - b. Thus,(a, b) = $(5a+2)u_1 + (2a - b)u_2$

(Step2) Now we find $F(u_1)$ and write it as a linear combination ofu_1 and u_2 and then we repeat the process for $F(u_2)$. We have; using the above formula for (a, b),

$F(u_1) = F(1, 2) = (4, 14) = 8u_1 - 6u_2$

$F(u_2) = F(2, 5) = (11, 33) = 11u_1 - 11u_2$

(Step3)

Finally, we write the coordinates of $F(u_1)$and $F(u_2)$as columns to obtain the required matrix:

$[F]_S = \begin{bmatrix} 8 & 11 \\ -6 & -11 \end{bmatrix}$

Our first theorem, tells us that the ‘‘action’’ of a linear operator T on a vector v is preserved by its matrix representation.

Theorem 12: let $T:V \longrightarrow$ □V, be a linear operator, and let S be a (finite) basis of V. Then, for any vector v in V, $[T]_S[v]_S = [T(v)]_S$.

Example 15 Consider the linear operator F on R^2 and the basis S of Example 6.3; that is,

$F(x, y) = (2x + 3y, 4x - 5y)$ and $S = \{u_1, u_2\} = \{(1, 2),(2,5)\}$

Let $v = (5, 7)$ and so $F(v) = (-11, 55)$

Using the formula from Example 6.3, we get

$[v]= [11, 3]^T$ and $[F(v)]= [55, -33]^T$

We verify Theorem 6.1 for this vector v (where [F]is obtained from Example 6.3):

$$[F][v] = \begin{bmatrix} 8 & 11 \\ -6 & -11 \end{bmatrix}\begin{bmatrix} 11 \\ -3 \end{bmatrix} = \begin{bmatrix} 55 \\ -33 \end{bmatrix} = [F(v)]$$

Given a basis S of a vector space V, we have associated a matrix [T]to each linear operator T in the operator algebra A(V) is preserved by this representation.

2.3 Change of Basis

Let V be an n-dimensional vector space over a field K. We have shown that once we have select a basis S of V, every vector v in V can be represented by means of an n-tuple $[v]_S$in K^n, and every linear operator T in A(V)can be represented by an $n \times n$ matrix over K. We first need a definition.

Definition: let $S = \{u_1, u_2,...,u_n\}$ be a basis of a vector space V, and let $S^0 = \{v_1,v_2, ..., v_n\}$ be another basis. (For reference, we will call S the ***old basis*** and S^0 the ***new basis***.) Since S is a basis, each vector in the ''new'' basis S^0 can be written uniquely as a linear combination of the vectors in S; say,

$$v_1 = a_{11} u_1 + a_{12}u_2 + ... + a_{1n}u_n$$
$$v_2 = a_{21} u_1 + a_{22}u_2 + ... + a_{2n}u_n$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$v_n = a_{n1} u_1 + a_{n2}u_2 + ... + a_{nn}u_n$$

Let P be the transpose of the above matrix of coefficients; that is, let $P = [p_{ij}]$, where $p_{ij}=a_{ji}$. Then P is called the change of basis matrix (or transition matrix) from the ''old'' basis S to the ''new'' basis S^0.

The following remarks are in order.

Remark 1: we may view the above change-of-basis matrix P as the matrix whose columns are, respectively, the coordinate column vectors of the ''new'' basis vectors v_i relative to the ''old'' basis S; namely,

$$P = [[v_1]_S, [v_2]_S, \ldots, [v_n]_S]$$

Remark 2: by way of analogy, there is a change-of-basis matrix Q from the ''new'' basis S^0 to the ''old'' basis S. Similarly, Q may be viewed as the matrix whose columns are the coordinate column vectors of the ''old'' basis vectors u_i, relative to the ''new'' basis S^0; namely,

$$Q = [[u_1]_S, [u_2]_S, \ldots, [u_n]_S]$$

Remark 3: Since the vectors v_1, v_2, ...,v_n in the new basis S^0 are linearly independent, the matrix P is invertible. Q is also invertible.

Example 16:Consider the following two bases of R^2:

$S = \{u_{1,} u_2\} = \{(1, 2), (3, 5)\}$and $\quad S^0 = \{\} = \{(1, 1), (1,2)\}$

(a) Find the change-of-basis matrix P from S to the ''new'' basis S^0.Write each of the new basis vectors of S^0 as a linear combination of the original basis vectors u_1 and u_2 of S. We have

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 3 \\ 5 \end{bmatrix} \text{or} \begin{matrix} x & + & 3y = & 1 \\ 2x & + & 5y = & -1 \end{matrix} \quad \text{yielding } x = -8 \text{ , } y = 3$$

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 3 \\ 5 \end{bmatrix} \text{or} \begin{matrix} x & + & 3y = & 1 \\ 2x & + & 5y = & -2 \end{matrix} \quad \text{yielding } x = -11 \text{ , } y = 4$$

Thus,

$v_1 = -8u_1 + 3u_2$

$v_2 = -11u_1 + 4_2u_2$and hence $P = \begin{bmatrix} -8 & -11 \\ 3 & 4 \end{bmatrix}$

Note that the coordinates of v_1 and v_2 are the columns, not rows, of the change of basis matrix P.

(a) Find the change of basis matrix Q from the ''new'' basis S^0 back to the ''old'' basis S.

Here we write each of the ‘‘old’’ basis vectors u_1 and u_2 of S^0 as a linear combination of the ‘‘new’’ basis vectors v_1 and v_2 of S^0. This yield

$u_1 = 4v_1 - 3v_2$

$u_2 = 11v_1 - 8v_2$ and hence $Q = \begin{bmatrix} 4 & 11 \\ -3 & -8 \end{bmatrix}$

As expected from Proposition 6.4, $Q = P^{-1}$. (In fact, we could have obtained Q by simply finding P^{-1}).

Applications of Change-of-Basis Matrix

First, we show how a change of basis affects the coordinates of a vector in a vector space V.

Theorem 13. Let P be the change-of-basis matrix from a basis S to a basis S^0 in a vector space V. Then, for any linear operator T on V,

$$[T]_{S^0} = P^{-1}[T]_S P$$

That is, if A and B are the matrix representations of T relative, respectively, to S and S^0, then

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

Example 17 Consider the following two bases of R^3:

$E = \{e_1, e_2, e_3\} = \{(1, 0, 0), (0,1,0), (0,0,1)\}$

And $S = \{u_1, u_2, u_3\} = \{(1, 0, 1), (2,1,2), (1,2,2)\}$

The change-of-basis matrix P from E to S and its inverse P^{-1} were obtained in Example 6.6.

(a) Write $v = (1, 3, 5)$ as a linear combination of u_1, u_2, u_3, or in equivalent, find $[V]_S$.

On the other hand, we know that $[V]_E = [1, 3, 5]^T$, because E is the usual basis and we already know P^{-1}, therefore,

$$[V]_S = P^{-1}[V]_E = \begin{bmatrix} -2 & -2 & 3 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 7 \\ -5 \\ 4 \end{bmatrix}$$

Thus, again $v = 7u_1 - 5u_2 + 4u_3$

(b) Let A=$\begin{bmatrix} 1 & 3 & -2 \\ 2 & -4 & 1 \\ 3 & -1 & 2 \end{bmatrix}$, which may be viewed as a linear operator on R . Find the matrix B that represents relative to the basis S.

The definition of the matrix representation of A relative to the basis of S tells us to write each of $A(u_1)$, $A(u_2)$, $A(u_3)$as a linear combination of the basis vectors u_1,u_2, u_3 of S. This yields

$A(u_1)$= (-1, 3, 5) = $11u_1 - 5u_2 + 6u_3$

$A(u_2)$= (1, 2, 9)=$21u_1 - 14u_2+8u_3$ and hence;

$$B = \begin{bmatrix} 11 & 21 & 17 \\ -5 & -14 & -8 \\ 6 & 8 & 2 \end{bmatrix}$$

$A(u_3)$= (3, -4, 5) =$17u – 8u_2 + 2u_3$

We emphasise that to find B, we need to solve three 3×3 systems of linear equations—one 3 × 3 system for each of $A(u_1)$, $A(u_2)$, $A(u_3)$.

On the other hand, because we know, P and P^{-1}, we can use Theorem 6.7. That is,

$$B = P^{-1}AP = \begin{bmatrix} -2 & -2 & 3 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 3 & -2 \\ 2 & -4 & 1 \\ 3 & -1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 11 & 21 & 17 \\ -5 & -14 & -8 \\ 6 & 8 & 2 \end{bmatrix}$$

This, as we expect, gives the same result.

2.4 Similar Matrices

Suppose A and B are square matrices for which there exists an invertible matrix P such that B =$P^{-1}AP$; then B is said to be similar to A, or B is said to be obtained from A by a similarity transformation.

Theorem 14: two matrices represent the same linear operator if the matrices are similar.

That is, all the matrix representations of a linear operator T form an equivalence class of similar matrices.

A linear operator T is *diagonal* if there exists a basis S of V such that we can represent T by a diagonal matrix; we can then say that the basis S is *diagonal to* T.

The preceding theorem gives us the following result.

Theorem 15. Let A be the matrix representation of a linear operator T. Then T is diagonal if there exists an invertible matrix P such that P^1AP is a diagonal matrix.

That is, T is diagonal if its matrix representation can be diagonally realised by a similarity transformation.

We emphasise that not every operator is diagonally realisable. However, we will show that every linear operator can be represented by certain ‘‘standard’’ matrices called its normal or canonical forms. Such a discussion will require some theory of fields, polynomials, and determinants.

Functions and Similar Matrices

Suppose f is a function on square matrices that assign the same value to similar matrices; that is, f (A) = f (B)whenever A is similar to B. Then f induces a function, also denoted by f on linear operator T, in the following natural way. We define

$$f(T) = f([T]_S)$$

Where S is any basis, the function is well defined.

The determinant is perhaps the most important example of such a function. The trace is another important example of such a function.

Example 18. Consider the following linear operator F and bases E and S of R^2:

$F(x;y) = (2x + 3y, 4x - 5y)$, $E = \{(1, 0), (0, 1)\}$, $S = \{(1, 2), (2, 5)\}$

The matrix representations of F relative to the bases E and S are, respectively,

$$A = \begin{bmatrix} 2 & 3 \\ 4 & -5 \end{bmatrix} \text{ and } B = \begin{bmatrix} 52 & 129 \\ -22 & -55 \end{bmatrix}$$

Using matrix A, we have

(h) Determinant of F = det(A) =10 -12 =-22;

(i) Trace of F =tr(A) = 2 -5 =-3:

On the other hand, using matrix B, we have

(j) Determinant of F =det(B)=-2860 +2838=-22;

(k) Trace of F = tr(B) =52 -55 =-3.

As expected, both matrices yield the same result.

3.0 Tutor Marked Assignment

1. Let F : $R^4 \rightarrow R^3$ be the linear mapping defined by

   F(x, y, z, t) = (x - y + z + t, 2x - 2y + 3z + 4t, 3x - 3y + 4z + 5t)

   (a) Find a basis and the dimension of the image of F.

   (b) Find a basis and the dimension of the kernel of the map F

2. The vectors u_1 = (1, 2, 0), u_2 = (1, 3, 2), u_3 = (0, 1, 3) form a basis S of R^3. Find

   (a) The change-of-basis matrix P from the usual basis E ={e_1, e_2, e_3}to S.

   (b) The change-of-basis matrix Q from S back to E.

3. Let $A = \begin{bmatrix} 4 & -2 \\ 3 & 6 \end{bmatrix}$ and $P = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

   (a) Find B = $P^{-1}AP$.

   (b) Verify tr(B) = tr(A)

   (c) Verify det(B) = det(A)

4.0 Conclusion/Summary

We have covered the following points in this study session.

1. Verify the linearity of certain mappings between vector spaces;
2. Construct linear transformations with certain specified properties;
3. Calculate the rank and nullity of a linear operator;
4. Define an isomorphism between two vector spaces;

5. Shown that two vector spaces are isomorphic if they have the same dimension;
6. Prove and use the Fundamental Theorem of homomorphism
7. Matrix representation of a linear map
8. Change of Basis
9. Similar matrix

5.0 Self-Assessment Questions

1. Consider the following linear operator T on R^2 and basis S:

$T(x, y) = (2x - 7y, 4x + 3y)$ and $S = \{u_1, u_2) = \{(1, 3), (2, 5)\}$.

Find the matrix representation $[T]_S$ of T relative to S.

2. Find the matrix representation of each of the following linear operators F on R^3 relative to the usual basis $E = \{e_1, e_2, e_3\}$ of R^3; that is, find $[F] = [F]_E$:

(a) F defined by $F(x, y, z) = (x + 2y - 3z, 4x - 5y - 6z, 7x + 8y + 9z)$.

(b) F defined by the 3×3 matrix $A=\begin{bmatrix}1 & 1 & 1\\ 2 & 3 & 3\\ 5 & 5 & 5\end{bmatrix}$

(c) F defined by $F(e_1) = (1, 3, 5)$, $F(e_2) = (2,4,6)$, $F(e_3) = (7,7,7)$.

Answer to Self-Assessment Questions:

1. $[T]_S = \begin{bmatrix}-121 & -201\\ 51 & 85\end{bmatrix}$

2. (a) $[F]_E = \begin{bmatrix}1 & 2 & -3\\ 4 & -5 & -6\\ 7 & 8 & 9\end{bmatrix}$

(b) $[F]_E = \begin{bmatrix}1 & 1 & 1\\ 2 & 3 & 3\\ 5 & 5 & 5\end{bmatrix}$

(c) $[F]_E = \begin{bmatrix}1 & 2 & 7\\ 3 & 4 & 7\\ 5 & 6 & 7\end{bmatrix}$

6.0 Additional Activities (Videos, Animations & Out of Class Activities)

Watch the videos:

a. http://www.patreon.com/patrickjmt

b. http://nptel.ac.in

c. http://ocw.mit.edu/18-06SCF11

7.0 References/Further Reading

Axler, S. et al (1997). Linear Algebra Done Right. 2nd Edition. New York: Springer-verlag

"MTH212 Linear Algebra II, National Open University of Nigeria, 2009 at www.nou.edu,ng

Lipschutz, S. (2009).Linear Algebra.4th Edition (Schaum's Outline Series). 2 Pennsylvania Plaza New York CityMcGraw-Hill

Study Session 2

Characteristic Polynomial and Characteristic Equation

Section and Subsection Headings

Introduction

In study session one (1), you have studied about the matrix of a linear transformation. You have had several opportunities, in earlier units, to observe that the matrix of a linear transformation depends on the choice of the bases of the concerned vector spaces.

Let V be an n-dimensional vector space over F, and let T: V⟶ V be a linear transformation. In this study session, we will consider the problem of finding a suitable basis B, of the vector space V, such that the n x n matrix [T] B is a diagonal matrix. This problem can also be seen as: given an n x n matrix A, find a suitable n x n non-singular matrix P such that P-1 AP is a diagonal matrix. It is in this context that the study of eigenvalues and eigenvectors play a central role.
The eigenvalue problem involves the evaluation of all the eigenvalues and eigenvectors of a linear transformation or a matrix.

The solution of this problem has basic applications in almost all branches of the sciences, technology and the social science beside its fundamental role in various branches of pure and applied mathematics. The emergence of computers and the availability of modern computing facilities has further strengthened this study, since they can handle very large systems of equations. We shall define eigenvalues and eigenvectors. We go on to discuss a method of obtaining them. In this session, we will also define the characteristic polynomial.

1.0 Study Session Learning Outcome

After studying this study session, you should be able to:

ii. Obtain the characteristic polynomial of a linear transformation or a matrix;
iii. Obtain the eigenvalues and eigenvectors of a linear transformation or a matrix;
iv. State and prove the Cayley-Hamilton theorem;

2.0 Main Content

2.1 Characteristic Polynomial and Characteristic Equations

Let A $= [a_{ij}]$ be an n-square matrix. The matrix M = A - tI_n, where $I_n$is the n-square identity matrix and t is an indeterminate, which you may obtain by subtracting t down the diagonal of A. The negative of M is the matrix tI_n- A, and its determinant

$$\Delta(t) = \det(tI_n - A) = (-1)^n \det(A - tI_n)$$
$$= t^n + a_1t^{n-1} + a_2t^{n-2} + \ldots + a_{n-1}t + a_n$$

Which is a polynomial in t of degree n and we call it *the **characteristic polynomial of A***.

Remark: suppose A $= [a_{ij}]$be a triangular matrix. Then tI_n– A is a triangular matrix with diagonal entries$t - a_{ij}$; hence,

$\Delta(t)$ = det $(t - a_{11})(t - a_{12}) \ldots (t - a_{1n})$

You will observe that the roots of D(t)are the diagonal elements of A.

Example 1 Let $A = \begin{bmatrix} 1 & 3 \\ 4 & 5 \end{bmatrix}$its characteristic polynomial is

$$\Delta(A) = |tI - A| = \begin{bmatrix} t-2 & -3 \\ -4 & t-5 \end{bmatrix} = (t-1)(t-5) - 12 = t^2 - 6t - 7$$

Characteristic Polynomials of Degrees 2 and 3

There are simple formulas for the characteristic polynomials of matrices of orders 2 and 3.

(a) Suppose $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Then

$$\Delta(t) = t^2 - (a_{11} + a_{22})t + \det(A) = t^2 - \mathrm{tr}(A)t - \det(A)$$

Here,tr(A)denotes the trace of A—that is, the sum of the diagonal elements of A.

(b) Suppose A = $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. Then

$$\Delta(t) = t^3 - \mathrm{tr}(A)\, t^2 + (A_{11} + A_{22} + A_{33})t - \det(A)$$

(Here, A_{11}, A_{22}, A_{33} denote, the cofactors of a_{11}, a_{22}, a_{33} respectively)

Example 2 Find the characteristic polynomial of $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 3 & 2 \\ 1 & 3 & 9 \end{bmatrix}$.

We have tr(A) =1 +3 +9 = 13. The cofactors of the diagonal elements are as follows:

$$A_{11} = \begin{bmatrix} 2 & 3 \\ 4 & -5 \end{bmatrix} = 21,\ A_{22} = \begin{bmatrix} 2 & 3 \\ 4 & -5 \end{bmatrix} 7, A_{33} = \begin{bmatrix} 2 & 3 \\ 4 & -5 \end{bmatrix} = 3$$

Thus, $A_{11} + A_{22} + A_{33} = 31$. Also, det(A) =27 +2 +0 -6 -6 -0 =17. Accordingly, $\Delta(t) = t^3 - 13t^2 + 31t - 17$

In-text Question 1

what is the highest power of the characteristic polynomial of an n × n matrices?

In-text Answer 1

The highest power/degree of the characteristic polynomial of an n × n matrix is n.

2.2 Diagonalisation, Eigenvalues and Eigenvectors

Let A be any n-square matrix. Then A can be represented by (or is similar to) a diagonal matrix D = diag($k_1, k_2, \ldots, k_n$) if and only if there exists a basis S consisting of (column) vectors $u_1, u_2, \ldots, u_n$ such that

$$Au_1 = k\, u_1$$
$$Au_2 = k_2\, u_2$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$Au_n = k_n\, u_n$$

In such a case, A can be ***diagonisable***. Furthermore, $D = P^{-1}AP$, where P is the non-singular matrix whose columns are the basis vectors $u_1, u_2 \ldots, u_n$ respectively.

The above observation leads us to the following definition.

Definition: Let A be any square matrix. A scalar λ is called ***an eigenvalue*** of *A* if there exists a nonzero (column) vector v, such that

$$Av = \lambda v$$

Any vector satisfying this relation is called ***an eigenvector*** *of A* belonging to the eigenvalueλ.

We note that each scalar multiple kvof an eigenvector v belonging to $\lambda$is also such an eigenvector, becauseA(kv) = k(Av) = k(λv) = λ (kv)

The set Eof all such eigenvectors is a subspace of V, which we call the ***eigenspace*** of λ. (If dim E_λ= 1, then$E_\lambda$is called an ***eigenline*** and $\lambda$is called a scaling factor.)

We sometimes use the terms characteristic value and characteristic vector (or proper value and proper vector), instead of eigenvalue and eigenvector.

The above observation and definitions give us the following theorem.

Theorem 16: An n-square matrix A is similar to a diagonal matrix D if A has n linearly independent eigenvectors. In this case, the diagonal elements of D,are the corresponding eigenvalues and D = $P^{-1}AP$, where P is the matrix whose columns are the eigenvectors.

Suppose a matrix A can be diagonalized as above, say $P^{-1}AP = D$, where D is diagonal. Then A has the extremely useful diagonal factorisation:

$$A = PDP^{-1}$$

Using this factorisation, the algebra of A reduces to the algebra of the diagonal matrix D, which we can calculate easily. Specifically, suppose D =diag(k_1, k_2, ..., k_n). Then

$$A^m = (PDP^1)^m = PD^mP^{-1} = P\,\mathrm{diag}(k_1{}^m, \ldots, k_n{}^m)P^{-1}$$

More generally, for any polynomial f (t),

$$f(A) = f(PDP^{-1}) = Pf(D)P^{-1} = P\,\mathrm{diag}(f(k_1), f(k_2), \ldots, f(k_n))\,P^{-1}$$

Furthermore, if the diagonal entries of D,are nonnegative, let

$$B= P\,\mathrm{diag}(p\sqrt{k_1}, \sqrt{k_2}, \ldots, \sqrt{k_n})\,P^{-1}$$

Then B is a nonnegative square root of A; that is, B^2 =A and the eigenvalues of B are nonnegative.

Example 3: Let A= $\begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$

(a) Find all eigenvalues and corresponding eigenvectors.

(b) Find a nonsingular matrix P such that D = $P^{-1}AP$ is diagonal, and P^{-1}.

Solution

(a) First find the characteristic polynomial $\Delta(t)$ of A:

$$\begin{aligned}\Delta(t) &= t^2 - tr(A)t + det(A) \\ &= t^2 - 5t + 4 \\ &= (t-1)(t-4)\end{aligned}$$

The roots $\lambda = 1$ and $\lambda = 4$ of $\Delta(t)$ are the eigenvalues of A. We find corresponding eigenvectors.

(i) Subtract $\lambda = 1$ down the diagonal of A to obtain the matrix M = A - λI, where the corresponding homogeneous system MX = 0 yields the eigenvectors to $\lambda = 1$. We have

$$M = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \text{ corresponding to } \begin{matrix} x + 2y = 0 \\ x + 2y = 0 \end{matrix} \text{ or } x + 2y = 0$$

The system has only one independent solution; for example, x = 2, y = -1. Thus, v_1 = (2, -1) is an eigenvector belonging to (and spanning the eigen space of) $\lambda = 1$.

(ii) Subtract $\lambda = 4$ down the diagonal of A to obtain

$$M = \begin{bmatrix} -2 & 2 \\ 1 & -1 \end{bmatrix} \text{ corresponding to } \begin{matrix} -2x + 2y = 0 \\ x - y = 0 \end{matrix} \text{ or } x - y = 0$$

The system has only one independent solution; for example, x = 1, y = 1. Thus, v_2 = (1, 1) is an eigenvector belonging to $\lambda = 4$

(b) Let P be the matrix whose columns are v_1 and v_2. Then

$$P = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \text{ and } D = P^{-1}AP = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \text{ where } P^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{-1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

Note: the characteristic polynomial $f_A(t)$ is a polynomial of degree n, it can have n roots at the most. Thus, an n x n matrix has n eigenvalues, at the most. For example, the matrix in Example 6 has two eigenvalues, 1 and –1, and the matrix in E5 has 3 eigenvalues

1) Eigenvectors of a linear transformation (or matrix) corresponding to distinct eigenvalues are linearly independent.

In-text Question 2
Does every eigenvalue correspond to a distinct eigenvector?

In-text Question 2
The statement is not necessarily true because some eigenvalues may not have eigenvector

2.3 Cayley-Hamilton Theorem

In this section, we present the **Cayley-Hamilton theorem**, which is related to the characteristic equation of a matrix. It is named after the British mathematicians Arthur Cayley (1821-1895) and William Hamilton (1805 – 1865, who were responsible for a lot of work done in the theorem of determinants.

Theorem 17 (Cayley-Hamilton): let $f(t) = t^n + c_1t^{n-1} + \ldots\ldots + c_{n-1}t + c_n$ be the characteristic Polynomial of an n x n matrix A. Then, $\Delta(t) = A^n + c_1 A^{n-1} + c_2A^{n-2} + \ldots. + c_{n-1}A + c_n1 = 0$(Note that over here 0 denotes the n x n zero matrix, and $I = I_n$.)

Proof: let A be an n-square matrix and let $\Delta(t)$ be its characteristic polynomial therefore

$$\Delta(t) = |tI - A| = t^n + a_{n-1} t^{n-1} + a_{n-2} t^{n-2} + \ldots. + a_1t + a_0$$

Let B(t) denote the adjoint of the matrix tI – A, the elements of B(t) are factors of tI – A and hence, are polynomials of degree not exceeding n – 1

Therefore,

$$B(t) = B_{n-1}t^{n-1} + B_{n-2}t^{n-2} + \ldots + B_1t + B_0$$

Where B_i is n-square matrices over K independent of t

But $(It - A)B(t) = |tI - A|.I$

i.e., $(It - A)(B_{n-1}t^{n-1} + B_{n-2}t^{n-2} + \ldots + B_1t + B_0) = (t^n + a_{n-1}t^{n-1} + a_{n-2}t^{n-2} + \ldots + a_1t + a_0)I$ removing the brackets and equating coefficient yields

$$
\begin{aligned}
B_{n-1} &= I \\
B_{n-2} - AB_{n-1} &= a_{n-1}I \\
B_{n-3} - AB_{n-2} &= a_{n-2}I \\
\vdots \quad & \quad \vdots \\
B_1 - AB_2 &= a_2 I \\
B_0 - AB_1 &= a_1 I \\
-B_0 &= a_0 I
\end{aligned}
$$

Multiplying the above equations by $A^n, A^{n-1}, A^{n-2}, \ldots, A^2, A, I$ respectively, we get -

$$
\begin{aligned}
A^n B_{n-1} &= A^n \\
A^{n-1}B_{n-2} - A^n B_{n-1} &= a_{n-1}A^{n-1} \\
A^{n-2}B_{n-3} - A^{n-1}B_{n-2} &= a_{n-2}A^{n-2} \\
\vdots \quad & \quad \vdots \\
A^2B_1 - A^3B_2 &= a_2A^2 \\
AB_0 - A^2B_1 &= a_1A \\
-AB_0 &= a_0
\end{aligned}
$$

Adding these equations, we have,

Thus, $\Delta(t) = A^n + a_{n-1}A^{n-1} + a_{n-2}A^{n-2} \ldots + a_1A + a_0I = 0$, and the Cayley-Hamilton theorem is proved.

Note: this theorem can also be stated as:

“Every square matrix satisfies its characteristic polynomial” or, “Every square matrix is a zero of its characteristic polynomial”

Example 4: Verify the Cayley-Hamilton theorem for $A = \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}$

Solution: The characteristic polynomial of A is

$$\begin{bmatrix} t-2 & -1 \\ 1 & t \end{bmatrix} = t^2 \ -3t + 2I$$

Therefore, we want to verify that A^2 -3A + 21 = 0.

Now, $A^2 = \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 7 & 6 \\ -3 & -2 \end{bmatrix}$

Therefore, $A^2 - 3A + 21 = \begin{bmatrix} 7 & 6 \\ -3 & -2 \end{bmatrix} - \begin{bmatrix} 9 & 6 \\ -3 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = 0$

Therefore, the Cayley-Hamilton theorem is true in this case.

3.0 Tutor Marked Assignment

1. Find the characteristic polynomials of each of the following matrices:

   (a) $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 3 \\ 5 & 5 & 5 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 3 \\ 5 & 5 & 5 \end{bmatrix}$

2. Let $A = \begin{bmatrix} 4 & 1 & -1 \\ 2 & 5 & -2 \\ 1 & 1 & 2 \end{bmatrix}$
   (a) Find all eigenvalues of A.
   (b) Find a maximum set S of linearly independent eigenvectors of A.
   (c) Is Adiagonalisable? If yes, find P such that $D = P^{-1}AP$ is diagonal.

4.0 Conclusion/Summary

We have covered the following point here.

i. The definition of eigenvalues, eigenvectors and eigenspaces of linear transformations and matrices.
ii. The definition of the characteristic polynomial and characteristic equation of a linear transformation (or matrix).

iii. A scalar λ is an eigenvalue of a linear transformation T (or matrix A) if it is a root of the characteristic polynomial of T (or A).

iv. A method of containing all the eigenvalues and eigenvectors of a linear transformation (or matrix).

v. A linear transformation T: V $\rightarrow$ V is diagonalizable ifV has a basis consisting of eigenvectors of t.

vi. Statement and the proof of Cayley-Hamilton theorem

5.0 Self-Assessment Questions

1. For each of the following matrices, find all eigenvalues and corresponding linearly independent eigenvectors:

(a) $A=\begin{bmatrix} 2 & -3 \\ 2 & -5 \end{bmatrix}$, (b) $B = \begin{bmatrix} 2 & 4 \\ -1 & 6 \end{bmatrix}$, (c) $C = \begin{bmatrix} 1 & -4 \\ 4 & -7 \end{bmatrix}$

When possible, find the nonsingular matrix P that diagonalises the matrix.

2. Verify the Cayley-Hamilton theorem for the matirices

$$A = \begin{bmatrix} 7 & 6 & 0 \\ 2 & 3 & 0 \\ -2 & -2 & 1 \end{bmatrix} \text{ and } B= \begin{bmatrix} 0 & 1 & 0 \\ 3 & 0 & 1 \\ 1 & -2 & -1 \end{bmatrix}.$$

6.0 Additional Activities (Videos, Animations & Out of Class Activities)

Watch the videos:

a. http://www.yuotube.com/watch?v=idsV0RaC9jM

b. http://www.MathResource.com

c. http://bit.ly/izBPlvm

7.0 References/Further Reading

Axler, S. et al (1997). Linear Algebra Done Right. 2nd Edition. New York: Springer-verlag

"MTH212 Linear Algebra II, National Open University of Nigeria, 2009 at www.nou.edu,ng

Lipschutz, S. (2009).Linear Algebra.4th Edition (Schaum's Outline Series). 2 Pennsylvania Plaza New York City McGraw-Hill

Glossary

Linear Equation

An equation of the form a1x1 + a2x2 + . . . + anxn = b where the aj and b stand for given quantities (e.g., numbers) and the xj are variables (unknowns). Occurrence of terms like x^2, i, x_1 x_2, cos x_j , $1/x_j$ would make the equation nonlinear.

Compatibility:

Two matrices are compatible (for multiplication) if the number of columns in the first equals the number of rows in the second. This is because each row (with *j* elements) in the first matrix will be multiplied by each column (with *i* elements) in the second matrix. For the element by element principle to work, the number of elements (columns) in the first must equal the number of elements (rows) in the second. Heuristic: m x n * o x p leads to NO GO if n < > o. Do they match where they touch?

Compatibility: the "inside" terms (c1 must equal r2)

$$\underbrace{(r_1 * \overbrace{c_1) * (r_2}^{compatibility} * c_2)}_{result}$$

Dimensionality of result : the "outside" terms (r1 * c2)

Determinant (Det):

<u>Square matrices only</u>. Det **A** = · aij (-1)i+j Det **M**ij, where **M**ij is the minor of **A** with the ith row and jth column deleted. (Sum over the j's for expanding by columns, which is the normal way to do it). For larger matrices, calculating the determinant is done by row interchanges. The determinant can tell us about the

number of solutions to a set of equations (unique, infinite?) and various other useful properties of the matrix.

Eigenvalues:

The characteristic roots of a set of simultaneous equations. In matrix form the eigenvalues (l) are defined such that Det(**A** - l**I**) = 0 [the characteristic equation]. The characteristic equation will be a polynomial in l of degree *n*, where *n* is the order (size) of the matrix. In demography, **the dominant eigenvalue is the population growth rate, l**, while the second eigenvalue gives the damping ratio (how quickly a population will return to equilibrium following a perturbation).

Eigenvector:

A vector, **u**, such that **A*u** = l***u.** For matrix models, the right and left eigenvectors corresponding to the dominant eigenvalue are the **reproductive values** (left eigenvector, a row vector) and the **stable (st)age distribution** (right eigenvector, a column vector).

Identity matrix:

Matrix usually written as **I**, with 1 (ones) on the main diagonal and zeros elsewhere.

Inverse:

$\mathbf{A}^{-1} * \mathbf{A} = \mathbf{I}$. The inverse of a matrix is that matrix which when premultiplied against **A**, yields the identity matrix. Inverse exists only when the matrix is singular (that is, has a determinant).

Irreducible:

Equivalent to "strongly connected" in diagraph theory (see Keyfitz p. 30). Each point on the graph (or in the matrix) can be reached from every other point, either directly or by going through intermediate nodes.

Leslie matrix:

Age-classified matrix used in life history analysis. It has elements only in the top row (fertility) and along the subdiagonal (survival). In order to produce a stable age distribution, the matrix must be **irreducible** (strongly connected) and **primitive**.

Matrix

A rectangular array of numbers (or symbols standing for numbers). An $m \times n$ matrix is a matrix with m rows and n columns. We usually denote matrices with capital letters like A, and their entries with the corresponding lowercase letters:

a_{ij} is the entry in the i^{th} row and j^{th} column of matrix A. Synonym: $a_{ij} = (A)_{ij}$ .The first index refers to the Row, the second index refers to the column.

Multiplication:

Done row by column, element by element. The product of r x c * c x s matrix is a matrix of order (= size) r x s (the number of rows in the first by the number of columns in the second). So, compatibility involves the inner two numbers and the 'order' of the product matrix involves the outer two numbers.

Primitive:

Any matrix for which $\mathbf{A}^n$ (for some arbitrary *n*) has only positive elements is said to be primitive. Primitivity of a nonnegative, irreducible matrix is a sufficient condition for stability. For the simple Leslie matrix this means that at least two age classes must exist that have *m*i > 0 and that have *i* relatively prime (e.g., for a six-year lifespan fecundity only at age-classes 3 and 6 will yield cycles rather than a stable age distribution).

Row and column vectors

A row vector is a $1 \times n$ matrix (a matrix with only one row), a column vector is a $m \times 1$ matrix (a matrix with only one column).

Singular:

A matrix whose determinant is zero.

Subscript notation:

For a matrix **A**, *a*ij means the element in the ith row and jth column. **A** = {aij}, where {} means "set of".

Symmetric:

A matrix is symmetric if it equals its own transpose. That is, if **A** + **A**'.

Trace:

Square matrices only. The sum of the elements of the main diagonal.

Transpose:

For a matrix **A**, the transpose **A**' = {arc}' = {acr} (use i and j as subscripts, if you prefer). That is, the elements in the rows of the first become the elements of the columns of the transpose.