ALGEBRA I (FOR NON-MAJOR)

MAT 213

MAT 213 - ALGEBRA I
(FOR NON-MAJOR)

by

Dr. M. EniOluwafe
Department of Mathematics
University of Ibadan,
Ibadan.

Ibadan Distance Learning Centre Series

MAT 213
Algebra I

By
M. EniOluwafe Ph.D.
Department of Mathematics
University of Ibadan

Published by
Distance Learning Centre
University of Ibadan

First Published 2008

ISBN 978-021-356-2

General Editor: Prof. Francis Egbokhare
Series Editors: Olubunmi I. Adeyemo and 'K. Ogunsola

Typeset @ Distance Learning Centre, University of Ibadan

Table of Content

Table of Content -Continue

Vice-Chancellor's Message

I congratulate you on being part of the historic evolution of our Centre for External Studies into a Distance Learning Centre. The reinvigorated Centre, is building on a solid tradition of nearly twenty years of service to the Nigerian community in providing higher education to those who had hitherto been unable to benefit from it.

Distance Learning requires an environment in which learners themselves actively participate in constructing their own knowledge. They need to be able to access and interpret existing knowledge and in the process, become autonomous learners.

Consequently, our major goal is to provide full multi media mode of teaching/learning in which you will use not only print but also video, audio and electronic learning materials.

To this end, we have run two intensive workshops to produce a fresh batch of course materials in order to increase substantially the number of texts available to you. The authors made great efforts to include the latest information, knowledge and skills in the different disciplines and ensure that the materials are user-friendly. It is our hope that you will put them to the best use.

Professor Olufemi A. Bamiro, FNSE
Vice-Chancellor

Foreword

The University of Ibadan Distance Learning Programme has a vision of providing lifelong education for Nigerian citizens who for a variety of reasons have opted for the Distance Learning mode. In this way, it aims at democratizing education by ensuring access and equity.

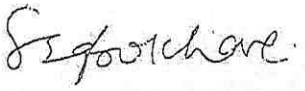
The U.I. experience in Distance Learning dates back to 1988 when the Centre for External Studies was established to cater mainly for upgrading the knowledge and skills of NCE teachers to a Bachelors degree in Education. Since then, it has gathered considerable experience in preparing and producing course materials for its programmes. The recent expansion of the programme to cover Agriculture and the need to review the existing materials have necessitated an accelerated process of course materials production. To this end, one major workshop was held in December 2006 which have resulted in a substantial increase in the number of course materials. The writing of the courses by a team of experts and rigorous peer review have ensured the maintenance of the University's high standards. The approach is not only to emphasize cognitive knowledge but also skills and humane values which are at the core of education, even in an ICT age.

The materials have had the input of experienced editors and illustrators who have ensured that they are accurate, current and learner friendly. They are specially written with distance learners in mind, since such people can often feel isolated from the community of learners. Adequate supplementary reading materials as well as other information sources are suggested in the course materials.

The Distance Learning Centre also envisages that regular students of tertiary institutions in Nigeria who are faced with a dearth of high quality textbooks will find these books very useful. We are therefore delighted to present these new titles to both our Distance Learning students and the University's regular students. We are confident that the books will be an invaluable resource to them.

We would like to thank all our authors, reviewers and production staff for the high quality of work.

Best wishes.



Professor Francis O. Egbokhare
Director

General Introduction and Course Objectives

The purpose of this course was to give a general introduction to algebra, building on what the students have learned from a standard first-year undergraduate course in algebra.

The course covers Sets, Relations and Partitions, Mappings, Matrix and System of Linear Equations, Determinant and Inverse Matrix, Vector spaces, Spanning Sets and Linear Independence, Homogeneous system, Linear transformations and their representation by matrices, Coordinates and Matrices under change of basis, Eigenvalues and Eigenvectors, Diagonalisation, Groups, Boolean Algebra, and Rings and Fields.

The objectives of this course are as follows. The reader should be able to

(1) exhibit a proficiency in the topics covered in the course.

(2) engage in critical thinking and problem solving; and

(3) translate technical information into mathematical statements, analyze information, formulate appropriate mathematical statements, solve problems and interpret solutions.

LECTURE 1

Sets

Introduction

We shall expose the meaning of a set and state the different ways of naming a set. We shall also reveal the different types of sets, subsets, equality of sets and the universe of discourse. We shall then interact various sets which belong to the same universe, using the definitions of union, intersection, power set, complements, relative complements and symmetric difference to form new sets.

Geometry representation of sets shall be presented in the form of Venn diagrams which will then be used in solving problems on sets. We also study some similarities of algebra of numbers to algebra of set theory and give a theorem on the number of elements in sets.

Objectives
At the end of this lecture you should be able to:
(i) have a working knowledge of the language of sets,
(ii) represent sets in Venn diagrams with a view to using them to solve problems on sets,
(iii) use laws of algebra of sets to simplify problems on sets, and
(iv) apply the understanding gained for Boolean algebra in lecture 14.

Pre-Test

See Post-Test at the end of the lecture.

Definition (Set)

A set is any well defined collection of "objects", known as the elements or members of the set.

Remark

The set must be well defined, so that there is no doubt which elements are included in the set, or excluded from it.

There are different ways of naming a set:

(i) the elements of the set may be enumerated by being shown within braces. For example

$$\{a, b, c, d\},$$

(ii) a set can be described by a sentence, for example

"all the people in this compound whose birthdays are in April",

(iii) a formal statement can be given, for example

$$E = \{x : x = 2n - 1;\ n \text{ is a natural number}\}$$

is read: E is the set of x's such that $x = 2n - 1$, and n is a natural number".

Remark

The set described formally under (iii) could be described by the sentence: "E is the set of positive odd numbers", or be enumerated as

$$E = \{1, 3, 5, 7, \cdots\}$$

The symbol $\in$ is used to say that certain elements are in the set.
Thus, if F is a set and f is an element of F, then we write $f \in F$. If an element g is not in F, then we write $g \notin F$.
For example, $5 \in E$ as above, but $8 \notin E$.

Definition and Notation (Subsets)

A subset is a set which contains some of the elements of another set. The symbol used is $\subset$, which is read either as "is a subset of" or "belongs to". On the other hand, the symbol $\not\subset$ stands for "is not a subset of" or "does not belong to". The symbol $\supset$ stands for "contains". Thus, $A \subset B$ if and only if $B \supset A$.

Fact

The empty set (is a set which contains no element), denoted by ϕ, is a subset of every set, by definition.

Definition (Proper and Improper Subsets)

A set is said to be an improper subset of itself. All other subsets are proper subsets, and ϕ is a proper subsets of all sets.

Standard Notation

We have the following for particular subsets of the number system:
$\mathbb{N} = \{0, 1, 2, \dots\}$ for the set of natural numbers;
$\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \dots\}$ for the set of integers;
$\mathbb{Q} = \{a/b : a, b \in \mathbb{Z}, \ \ b \neq 0\}$ for the set of rationals,
$\mathbb{R}$ for the set of real numbers;
$\mathbb{C}$ for the set of complex numbers;
$(a, b] = \{x \in \mathbb{R} : \ \ a < x \leq b\}$,
$(a, b) = \{x \in \mathbb{R} : \ \ a < x < b\}$, etc.

Definition and Notation (The universe of discourse)

The universe of discourse, denoted by $\mathcal{U}$, is the larger set of which other sets are parts.
For example, within

$$\mathcal{U} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\},$$

we can have $H = \{5, 7, 9, 11\}$ and $K = \{4, 5, 6, 7, 9\}$.

Note

The universe of discourse is not unique as it changes from problem to problem.

Fact

Every set is a subset of its universal set $\mathcal{U}$.

Definition (Equality of Sets)

A set is equal to another set only if both sets have the same elements.

Remark

It is not necessary for the elements in the two sets to be arranged in the same order. Also, repeated elements are ignored. Thus, if a set F has elements $8, 6, 7, 5$, and a set G has elements $5, 8, 7, 6, 6, 5$, we can say that $F = G$.
Two (or more) sets can interact if they belong to the same universe $\mathcal{U}$. The most obvious comparison is to seek the elements which two sets have in common.

Definition (Intersections of Sets)

Let S and T be any two sets. Then we define the intersection, denoted by $S \cap T$ of S and T as the set $S \cap T = \{x : x \in S \text{ and } x \in T\}$. Intersection has the symbol $\cap$, which is read either as "intersection" or "cap".
For example,
Let $S = \{1, 2, 5, 6, 7\}$ and $T = \{1, 2, 3, 4, 7\}$, then $S \cap T = \{1, 2, 7\}$.

It may happen that two sets have no common elements. In such a case the sets are said to be disjoint, and their intersection is empty.
For example, within

$$\mathcal{U} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\},$$

we can have $P = \{1, 7\}$ and $Q = \{5, 6\}$, then $P \cap Q = \phi$.

Note

If $X = \{0, 1, 2\}$ and $Y = \{0, 7, 8\}$ then $X \cap Y = \{0\}$. Now, since 0 is an element of $\mathcal{U}$, $\{0\}$ is not the same as ϕ. However, ϕ is sometimes written $\{\ \}$. Another operation on two sets is to ask for the set of elements that are in both sets.

Definition (Union of Sets)

Let S and T be any two sets. Then we define the union, denoted $S \cup T$ of S and T as the set

$$S \cup T = \{x : x \in S \text{ or } x \in T\}.$$

Union has the symbol $\cup$, which is read either as "union" or "cup".
In the examples above,

$$\begin{aligned} S \cup T &= \{1, 2, 3, 4, 5, 6, 7\}; \\ P \cup Q &= \{1, .5, 6, 7\}; \\ X \cup Y &= \{0, 1, 2, 7, 8\}. \end{aligned}$$

An example that illustrates relationship between subset, union and intersection,
$P = \{1, 7\}$ is a subset of $S = \{1, 2, 5, 6, 7\}$, or
$T = \{1, 2, 3, 4, 7\}$ of $S \cap T = \{1, 2, 7\}$, of
$S \cup T = \{1, 2, 3, 4, 5, 6, 7\}$, as well as of
$\mathcal{U} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.
We can write $P \subset S$, $P \subset (S \cap T)$, etc. It is also true to write

$$S \cap P = P \quad \text{and} \quad S \cup P = S.$$

Verify this.

Definition and Notation (Power set of a set)

This is the set of all the subsets of the set. For example, the set $A = \{1, 2, 3\}$ has an improper subset $\{1, 2, 3\}$ and the seven proper subsets $\{1, 2\}$, $\{1, 3\}$ $\{2, 3\}$, $\{1\}$, $\{2\}$, $\{3\}$ and ϕ.

The set which has these eight sets as its elements, is called the power set of A and is denoted by $\mathcal{P}(A)$.
For finite sets, the power set of A is always bigger than the set itself. In fact,

$$|A| = n \Rightarrow |\mathcal{P}(A)| = 2^n.$$

Practice Exercise

List the subsets that form the power set of $\{R, S, T, U\}$.

Definitions and Notations (Complements and relative complements)

A complement is a set of elements contained in the universe $\mathcal{U}$, but not in a particular subset of it.
For $S = \{1, 2, 5, 6, 7\}$, its complement
not-S (denoted S') in $\mathcal{U} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ is $S' = \{0, 3, 4, 8, 9\}$; similarly, $T' = \{0, 5, 6, 8, 9\}$.
Also, $S' \cap T' = \{0, 8, 9\}$.

Practice Exercise

Enumerate for the sets $P = \{1, 7\}$ and $X = \{0, 1, 2\}$ within the universe $\mathcal{U} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, the sets $P', X', P' \cup X'$ and$(P \cup X)'$.
Notice the notation and the results, make Venn diagrams if you have difficulty.

Definitions and Notation (relative complement)

The relative complement is a comparison between two sets. Not S in T means the set of elements of T which are not also in S. This relative complement, symbolized as $T - S$ is $\{3, 4\}$. Not - T in S is $S - T = \{5, 6\}$, where $T = \{1, 2, 3, 4, 7\}$ and $S = \{1, 2, , 5, 6, 7\}$.

Practice Exercises

1. Make a similar diagram to show S' and $S - T$. Also find out whether you agree that $S - T = S \cap T'$ and $T - S = T \cap S'$.

2. Draw a Venn diagram of three intersecting sets, A, B, C . There are eight regions including the part of $\mathcal{U}$ outside A, B, C. Describe these regions in terms of the symbols A, B, C, A', B', C', and $\cap$. For example, the middle region is $A \cap B \cap C$ and the outer part of A is $A \cap B' \cap C'$.

3. Experiment with these ideas and notations. What special results do you get if:

   (a) the sets T and S are disjoint;

   (b) S is contained within T?

4. Satisfy yourself that $B' \cap C' = (B \cup C)'$, and find similar relationships for A' and B', and for A' and C'.

Definition and Notation (symmetric difference)

The symmetric difference set $(A \cap B') \cup (A' \cap B)$ will be denoted by $A \triangle B$.

Practice Exercise

Give your example to verify the symmetric difference set as defined above.

Venn Diagrams

By drawing two or more closed shapes (not necessarily circles) that overlap, it is possible to illustrate definitions and operations on sets. We represent a universal set by a rectangle while we use circles to represent subsets of this universal set as shown in Fig. 1.1 and Fig. 1.2.

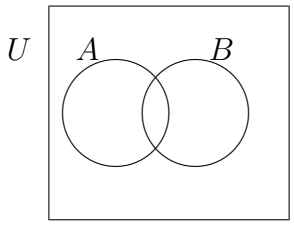


Fig 1.1

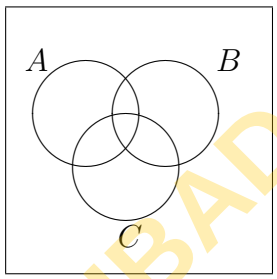


Fig. 1.2

Thus Figs 1.1 and 1.2 are Venn diagrams with two and three subsets of a universal set, respectively.

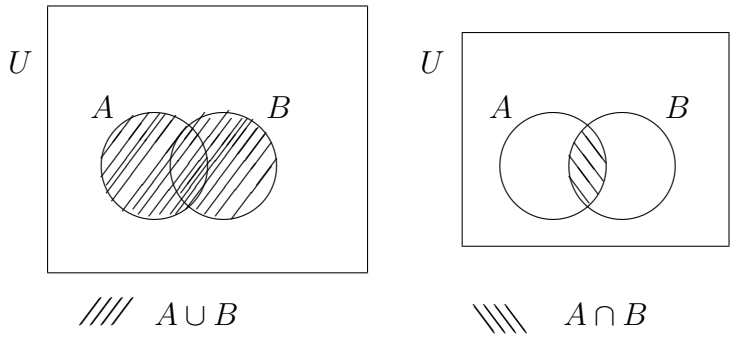


Fig 1.3

Fig. 1.3 shows the Venn diagrams of union and intersection.
.

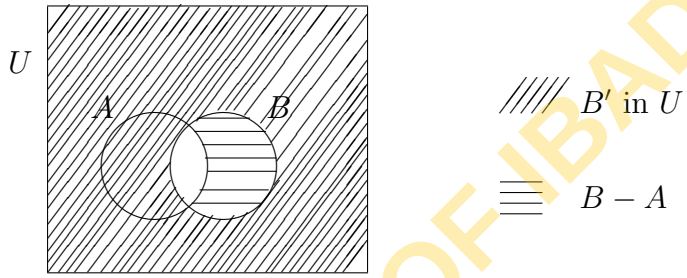


Fig 1.4

Fig. 1.4 shows Venn diagrams of the difference between B' and $B - A$.

The algebra of set theory

We show a table below, illustrating certain similar laws or propositions as applied in algebra of numbers with respect to operations $+$ and $\times$ to operations union and intersection in algebra of sets, respectively.

Table 1.1

Law	Numbers: a,b,c Operations: $+$, $\cdot$	Sets: A,B,C Operations: $\cup, \cap$
Commutative	$a+b=b+a$ $a \cdot b = b \cdot a$	$A \cup B = B \cup A$ $A \cap B = B \cap A$
Associative	$a+(b+c)=(a+b)+c$ $a \cdot (b \cdot c) = (a \cdot b) \cdot c$	$A \cup (B \cup C) = (A \cup B) \cup C$ $A \cap (B \cap C) = (A \cap B) \cap C$
Distributive	$a \cdot (b+c) = a \cdot b + a \cdot c$	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Illustrations by means of Venn diagrams, show whether or not these relationships are reasonable. It must be emphasized that an illustration is not a proof.

The last law is the most surprising, since it is not true for algebra:
For example

$$8 + (5 \times 3) \neq (8+5) \times (8+3)$$

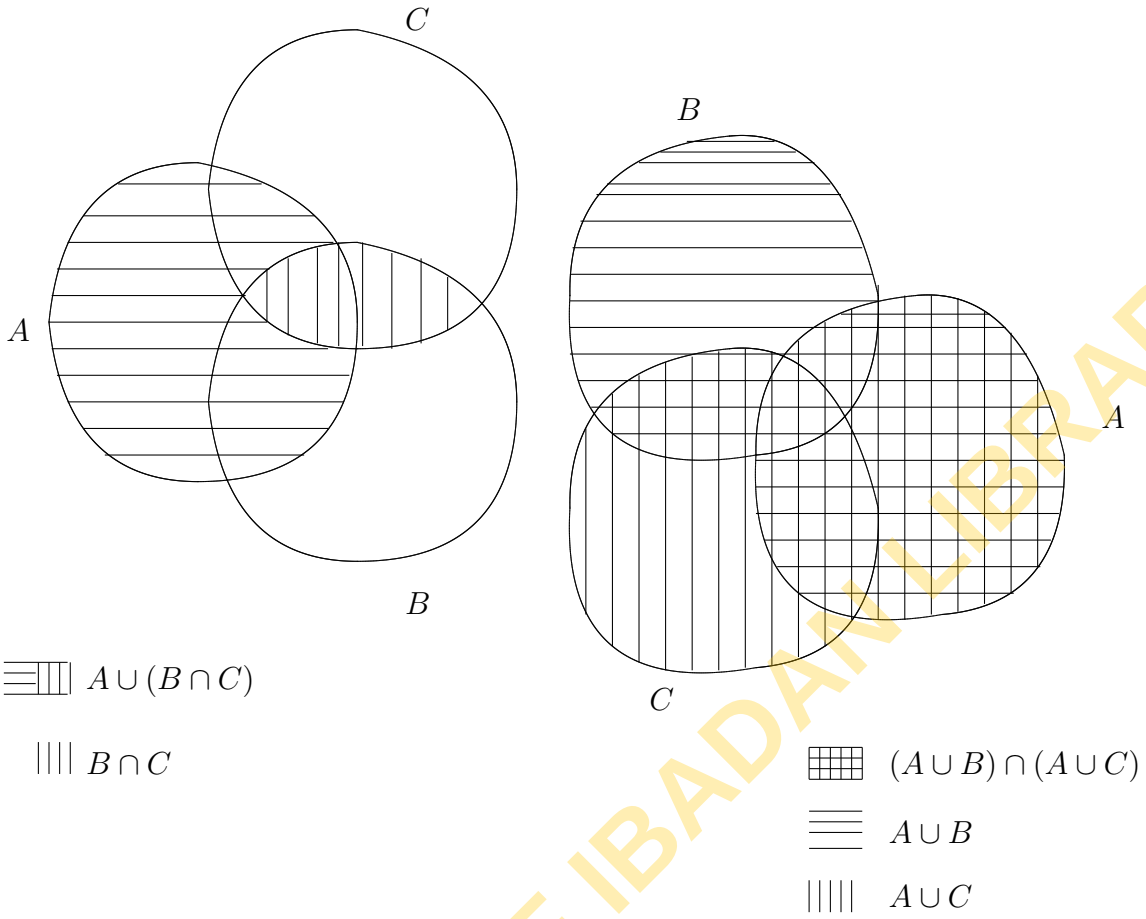


Fig. 1.5

Fig. 1.5 The second distributive law for sets

Fig. 1.5 reveals that it does make sense with sets.

Practice Exercise

Illustrate all the other rules for yourself, and also the following special ones: $\emptyset\,' = U,\ U' = \emptyset,\ A \cup A = A,\ A \cap A = A,\ A \cup A' = \emptyset,\ A \cup \emptyset = A,\ A \cap \emptyset = \emptyset,\ (A')' = A$;

De Morgan's laws:

$(A \cap B)' = A' \cup B'$; $(A \cup B)' = A' \cap B'$.

Number of Elements in a Set

Example: 1000 people were asked if they watched TV and listened to the radio and, if so, which they preferred. 130 said "neither", 720 said "TV" and 500 said "radio". How many of the 1000 both watched TV and listened to the radio?

A new use of Venn diagrams can help:

.

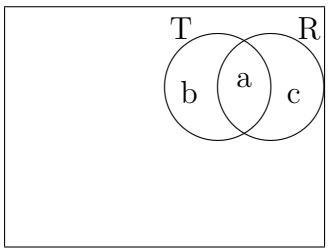


Fig. 1.6

$$\begin{array}{rrrr} T \cup R & b + a + c = & 870 \\ T & a + b & = & 720 \\ R & a + c & = & 590 \end{array}$$

On solving, $a = 440$.

Note that a, b, c stand for the number of elements.

Theorem(Number of Elements in a Set)

Call $n(T)$ or $[T]$ the number of elements in T, $n(R)$ the number of elements in R, $n(T \cup R)$ the number in $(T \cup R)$, $n(T \cap R)$ the number in $(T \cap R)$. Then $n(T \cup R) = n(T) + n(R) - n(T \cap R)$.

For example, since $870 = 720 + 590 - n(T \cap R)$, hence $n(T \cap R) = 440$.

Practice Exercises

1. Work out the same corresponding theorem for three sets.
2. Find or invent some problems of this kind and try out different ways of finding solutions.
3. How does the theorem read if T and R are disjoint sets?
4. A survey of 120000 families disclosed that 90000 had a car. 80000 had a TV set and 25000 had a boat. 10000 of the families had all three, and 60000 had both car and TV. 14000 had a car and a boat. If the same number of families had none of these things as had only a boat, calculate how many had only TV and how many had both a boat and TV but no car.

Summary
When sets have been defined, operations are introduced so that the sets can act on one another. An algebra of sets emerges and certain "laws" are obeyed; uses are found for this algebra. In particular, the understanding gained is needed for Boolean algebra in Lecture 14.

Post-Test

1. What remarks can be made about one set that is not the empty set?
2. Two sets A and B have some common elements. What operations, laws or problems on A and B can usefully be illustrated or solved with help from Venn diagrams?
3. $\mathcal{U} = \{1, 2, 3, 4, 5, 6\}$, $A = \{1, 2, 3\}$, $B = \{4, 5, 6\}$ What can you can about the sets A and B?

4. Let A, B, C be subsets of a set X. Simplify the expressions

   (a) $(A \cup (A \cup B)')'$;

   (b) $((A \cup \phi) \cap (B \cup A') \cap (A \cup B' \cup X))'$,

   (c) $(A \cup (B \cap C) \cup (B' \cap C') \cup C)'$.

5. Let A, B be subsets of a set X. Prove using a Venn diagram, that $(A \cup B') \cup (A' \cap B) = A \cup B \Leftrightarrow A \cap B = \phi$.

Supplementary Reading

1. A.,O. Kuku, Abstract Algebra, Ibadan University Press, 1980, Chapter 1, pp. 3-13.

2. S.A. Ilori and O. Akinyele, Elementary Abstract and Linear Algebra, Ibadan University Press, 1986. Chapter 1, pp 1-26.

3. S. Lipschutz, Set Theory and Related Topics, Schaum's Outline Series.

4. F. Ayres, Modern Algebra, Schaum's Outline Series.

5. S. Lipschutz, General Topology, Schaum's Outline Series Chapter 1.

6. G. Birkoff and S. MaClane, A survey of Modern Algebra, Macmillan Co. 1965.

7. J.B. Fraleigh, A First Course in Abstract Algebra, Addison Wesley 1968.

8. E.H. Connell, Elements of Abstract and Linear Algebra. (http://www.math.miami.edu/∼ec/book/)

LECTURE 2

Relations and Partitions

Introduction

Having dealt with operations on elements and sets, we now consider connections between sets as subsets of their Cartesian product. We shall use the properties of the subsets which is also called relations and consider some important classes of relations of which equivalence relations are probably the most important.

Objectives
At the end of this lecture you should be able to:
• distinguish properties of relations, and
• use an equivalence relation to partition the set on which it is defined.

Pre-Test

1. For each of the following relations, determine whether or not that the relation is reflexive, symmetric, transitive, an equivalence relation, giving appropriate reasons for your answers:

   (a) the relation P on the set $\mathbb{N}$ of natural numbers, where natural numbers m and n satisfy mPn if and only if $m+n$ is divisible by 2,

   (b) the relation Q on the set $\mathbb{N}$ of natural numbers, where natural numbers m and n satisfy mQn if and only if $m+n$ is divisible by 3,

   (c) the relation R on the set $\mathbb{N}$ of natural numbers, where natural numbers m and n satisfy mRn if and only if $n = 2^k m$ for some integer k (which may be positive, zero or negative),

(d) the relation S on the set $\mathbb{Z}$ of integers, where integers x and y satisfy xSy if and only if $x^2 < y^2$, and

(e) the relation Q on the set $\mathbb{Z}$ of integers, where integers x and y satisfy xQy if and only if $x - y = k^3$ for some integer k.

2. Let R be the relation on $\mathbb{N}$ defined by $aRb \Leftrightarrow a^2 \equiv b^2 \pmod 7$.
Show that R is an equivalence relation. Into how many equivalence classes does R partition $\mathbb{N}$?

3. Prove that the relation R defined on $\mathbb{R}$ by

$$xRy \Leftrightarrow x^2 - y^2 = 2(y - x)$$

is an equivalence relation. Determine the R-class of 0 and the R-class of 1.

4. For the given set and relations below, determine which define equivalence relations

(a) S is the set of all people in the world today,
$a \sim b$ if a and b have an ancestor in common.

(b) S is the set of all people in the world today,
$a \sim b$ if a and b have the same father.

(c) S is the set of real numbers $a \sim b$ if $a = \pm b$.

(d) S is the set of all straight lines in the plane,
$a \sim b$ if a is parallel to b.

2. Relations, Partitions

Definition: (Cartesian Products of Sets)
Let A and B be sets. The Cartesian product $A \times B$ of the sets A and B is defined to be the set of all ordered pairs (a, b) with $a \in A$ and $b \in B$.

Such an ordered pair (a, b) is comprised of two elements a and b, where the first element a is taken from the set A, and the second element b is taken

from the set B. If (a_1, b_1) and (a_2, b_2) are ordered pairs of this type then $(a_1, b_1) = (a_2, b_2)$ if and only if $a_1 = a_2$ and $b_1 = b_2$.

Example

Let $A = \{1, 2, 3\}$ and$B = \{1, 2\}$. Then

$$A \times B = \{(1, 1), (2, 1), (3, 1), (1, 2), (2, 2), (3, 2)\}.$$

Note that, in this example, the number of elements of the set $A \times B$ (i.e., 6) is the product of the number of elements of A (i.e., 3) and the number of elements of B (i.e., 2).

Example

Consider a set of children $P = \{a, b, c, d\}$ and the set of days in a particular week $Q = \{m, t_1, w, t_2, f, s_1, s_2\}$. Then

$$\begin{aligned} P \times Q \quad = \quad & \{(a, m), (b, m), (c, m), (a, t_1), (b, t_1), (c, t_1), \\ & (a, w), (b, w), (c, w), (a, t_2), (b, t_2), (c, t_2) \\ & (a, f), (b, f), (c, f), (a, s_1), (b, s_1), (c, s_1) \\ & (a, s_2), (b, s_2), (c, s_2)\} \end{aligned}$$

Example

Points of the plane are located in Cartesian coordinates by means of ordered pairs (x, y), where x and y are real numbers. The set of such ordered pairs is the set $\mathbb{R} \times \mathbb{R}$ (the Cartesian product of two copies of the set $\mathbb{R}$ of real numbers).

One may form the Cartesian product of any number of sets. Suppose that $A_1, A_2, \ldots, A_n$ are sets. Then the Cartesian product of these sets is the set $A_1 \times A_2 \times \cdots \times A_n$ consisting of all ordered n-tuples $(a_1, a_2, \ldots, a_n)$ with $a_i \in A_i$ for $i = 1, 2, \ldots, n$.

Example

Points of three dimensional space are specified in Cartesian coordinates by means of ordered triples (x, y, z), where x, y and z are real numbers. The set of such ordered triples is the set $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$.

A Cartesian product $A_1 \times A_2 \times \cdots \times A_n$ of finite sets $A_1, A_2, \ldots, A_n$ is itself a finite set: the number of elements of the Cartesian product is equal the

product of the number of elements of the individual sets $A_1, A_2, \ldots, A_n$.

Example

If the sets A, B and C have 4, 8 and 10 elements respectively then their Cartesian product has 320 elements, since $320 = 4 \times 8 \times 10$.

Example

Suppose that one is to construct a database containing information on students taking a course such as MAT 213. Each record in the database is to specify the student number, the name, and the degree programme being followed by the student. Let S be the set consisting of all strings of eight decimal digits. Let M be a set containing all the student names, and let N be the set of all degree programmes taught at University of Ibadan. Then a record in the database determines an element of the set $S \times M \times N$, such as

(63009987, Michael Oke Joel, BSc. Nigeria)

The collection of all such records contained in the database can be viewed as a subset of the Cartesian product $S \times M \times N$ of the sets S, M and N. The language of sets and Cartesian products is used in discussions of relational databases.

A subset of the Cartesian product $A_1 \times A_2 \times \cdots \times A_n$ of sets $A_1, A_2, \ldots, A_n$ is sometimes referred to as an n-ary relation on the sets $A_1, A_2, \ldots, A_n$.

Definition: (Binary Relation)

A binary relation is a subset of the Cartesian product $A_1 \times A_2$ of sets A_1 and A_2. Thus a binary relation on a set specifies relations between pairs of elements from the set.

Example

The relations $=$ ("equals"), $\neq$ ("not equal to"), $<$ ("less than"), $>$ ("greater than"), $\leq$ ("less than or equal to") and $\geq$ ("greater than or equal to") are all binary relations on the set $\mathbb{R}$ of real numbers.

Example

Let A be a set, and let $\mathcal{P}(A)$ be the power set of A (i.e., the set whose elements are the subsets of A). Then $\subset$ is a binary relation on $\mathcal{P}(A)$, where two subsets B and C of A satisfy $B \subset C$ if and only if B is a subset of C.

If one has a relation R on a set A, then, given two elements x and y of A, either x is related to y, in which case we may write xRy, or else the element is not related to y.

Definition (Equivalence relation)

Let R be a relation on a set A.
The relation R is said to be reflexive when it has the following property:
xRx for all elements x of the set A.
The relation R is said to be symmetric when it has the following property:
if x and y are elements of the set A, and if xRy, then yRx.
The relation R is said to be transitive when it has the following property:
if x, y and z are elements of the set A, and if xRy and yRz, then xRz.

An equivalence relation is a relation that is reflexive, symmetric and transitive.

Example

The relation $=$ ("equals") on the set $\mathbb{R}$ of real numbers is an equivalence relation. However, none of the relations $\neq$ ("not equal to"), $<$ ("less than"), $>$ ("greater than"), $\leq$ ("less than or equal to") or $\geq$ ("greater than or equal to") are equivalence relations on $\mathbb{R}$.

Examples

The relations "has the same hair color as" or "is the same age as" in the set of people are equivalence relations.

Definition (Equivalence class)

Let $\sim$ be an equivalence relation on S. For $a \in S$, the set of all elements equivalent to a is denoted by

$$[a] = \{b \in S : b \sim a\}$$

and called the equivalence class of a.

Examples

The equivalence classes under the relation "has the same hair color as" are the set of blond people, the set of red-haired people, etc.

Definition (Partition)

A partition of a nonempty set S is a collection $\{A_1, \ldots, A_n\}$ of nonempty subsets of S, called the cells (or blocks) of the partition, for which

(1) $A_i \cap A_j = \phi$ for all $i \neq j$

(2) $S = A_1 \cup \cdots \cup A_n$.

Example

Consider the following set of Countries

$$= \{\text{Germany, England, India, China, USA, Canada, Nigeria, Ghana}\}$$

One possible partition on countries is one that classifies them according to the continents they are in. We get the following partition

$$\text{Part}_C := \{\{\text{Germany, England}\}, \{\text{India, China}\}, \{\text{USA, Canada}\}, \{\text{Nigeria, Ghana}\}\}.$$

(Note that a partition is always a set of sets).

Example

Let $\mathbb{Z}$ be the set of integers, let O be the set of odd integers, and let E be the set of even integers. Every integer is either even or odd, and no integer is both even and odd.

Therefore, any integer belongs to exactly one of the sets O and E. Thus, the collection consisting of the sets O and E is a partition of the set $\mathbb{Z}$ of integers.

The next result establishes a one-to-one correspondence between equivalence relations on S and partition of S.

Theorem(Fundamental Theorem of Equivalence Relations)

(1) Let $\sim$ be an equivalence relation on S. Then the set of distinct equivalence classes with respect to $\sim$ are the cells of a partition of S.

(2) Conversely, if $\mathcal{P}$ is a partition of S, the binary relation $\sim$ defined by $a \sim b$ if a and b lie in the same block of $\mathcal{P}$ is an equivalence relation on S, whose equivalence classes are the blocks of $\mathcal{P}$.

Summary

We define Cartesian product between sets and consider a relation as any subject of the Cartesian product of two sets. We then consider a special type of relation called an equivalence relation. Finally we state a theorem which establishes a one-to-one correspondence between equivalence relations on a given set and partitions of the given set.

Post-Test

See Pre-Test at the beginning of the lecture.

Supplementary Reading

1. A.,O. Kuku, Abstract Algebra, Ibadan University Press, 1980, pp. 14-24.
2. S.A. Ilori and O. Akinyele, Elementary Abstract and Linear Algebra, Ibadan University Press, 1986. pp. 130-133, 150-156.
3. S. Lipschutz, Set Theory and Related Topics, Schaum's Outline Series.
4. F. Ayres, Modern Algebra, Schaum's Outline Series.
5. S. Lipschutz, General Topology, Schaum's Outline Series Chapter 1.
6. G. Birkoff and S. MaClane, A survey of Modern Algebra, Macmillan Co. 1965.
7. J.B. Fraleigh, A First Course in Abstract Algebra, Addison Wesley 1968.

LECTURE 3

Mappings between Sets

Introduction

We now consider another relations called mappings. We shall also consider the various types of mappings.

Objectives
At the end of this lecture you should be able to define, identify and solve problems involving:
- one-to-one, onto and bijective mappings,
- identity and inverse mappings, and
- composition of mappings.

Pre-Test

1. For each of the following maps, determine whether or not that map is injective and/or surjective and whether or not it has a well-defined inverse, giving appropriate reasons for your answers.

   (a) the map $f : \{1, 2, 3, 4\} \longrightarrow \{1, 2, 3, 4\}$ with $f(1) = 2$, $f(2) = 3$, $f(3) = 2$ and $f(4) = 4$;

   (b) the map $g : \{1, 2, 3, 4\} \longrightarrow \{1, 2, 3, 4\}$ with $g(1) = 2$, $g(2) = 3$, $g(3) = 1$ and $g(4) = 4$;

   (c) the map $h : [1, 2] \longrightarrow [0, \frac{1}{2}]$ with

   $$h(x) = \frac{x - 1}{x},$$

   where $[1, 2] = \{x \in \mathbb{R} : 1 \leq x \leq 2\}$ and $[0, \frac{1}{2}] = \{x \in \mathbb{R} : 0 \leq x \leq \frac{1}{2}\}$]

(d) the map $f : [-1, 1] \longrightarrow [-2, 2]$ with $f(x) = x^3 - x$ for all $x \in [-1, 1]$

(e) the map $g : [1, 2] \longrightarrow [0, 6]$ with $g(x) = x^3 - x$ for all $x \in [1, 3]$.

(f) the map $h : [0, 1] \longrightarrow [-2, 2]$ with $h(x) = x^3 - x$ for all $x \in [0, 1]$.

(g) the map $f : [-1, 1] \longrightarrow [-2, 2]$ with $f(x) = x^3 + x$ for all $x \in [-1, 1]$.

(h) the map $g : (-1, 1) \longrightarrow \mathbb{R}$ with $g(x) = \dfrac{1}{1 - x^2}$ for all $x \in (-1, 1)$.

2. Find the domain and image of the following relations on $\mathbb{R}$. Are any of them mappings?

   (a) $\{(x, y)|x^2 + 4y^2 = 1\}$

   (b) $\{(x, y)|x^2 = y^2\}$

   (c) $\{(x, y)|y = 2x - 1\}$

3. Let $f, g : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = (x + 1)^2$, $g(x) = 2x - 1$. Determine the mapping $f \circ g$, $g \circ f$ and the set

$$\{x \in \mathbb{R}|(f \circ g)(x) = (g \circ f)(x)\}.$$

Definition (Domain and Codomain of a map)

Let A and B be sets. A map $f : A \to B$ from A and B assigns to each element a of A an element $f(a)$ of B. The set A on which the map is defined is referred to as the domain of the map $f : A \to B$. The set B into which the domain is mapped by is referred to as the codomain of the map f.

Example

Let $\mathbb{R}$ be the set of real numbers. The map $q : \mathbb{R} \to \mathbb{R}$ defined by $q(x) = x^2$ for all real numbers x is a function from the set $\mathbb{R}$ of real numbers to itself.

Example There is a map $f : \mathbb{R}\backslash\{0\} \longrightarrow \mathbb{R}$, where $f(x) = 1/x$ for all non-zero real numbers x. The domain of this map is the set $\mathbb{R}\backslash\{0\}$ of all non-zero real numbers (i.e., the set $\{x \in \mathbb{R} : x \neq 0\}$). The domain of this map cannot be extended to the entire set $\mathbb{R}$ of real numbers since the reciprocal of zero is not defined. According to the above definition the value of a map must be defined at all elements of its domain.

Given any set A, there is a map $1_A : A \to A$ from the set A to itself which sends each element a of A to itself. The map is referred to as the identity map on A.

Definition (Range of a map)

Let A and B be sets, and let $f : A \to B$ be a map from A to B. The range of the map f is the subset $f(A)$ of B defined by

$$f(A) = \{b \in B : b = f(a) \text{ for some } a \in A\}.$$

In other words, the range of a map is the set consisting of all elements of the codomain of the map that are images under the map of elements of its domain.

Definition: (Compositions of Mappings)

Let A, B and C be sets, let $f : A \to B$ be a map A to B, let $g : B \to C$ be a map from B to C. Then there is a map $g \circ f : A \to C$ obtained by composing the maps f and g. This map is defined at each element a of A by the formula $(g \circ f)(a) = g(f(a))$. (In other words, in order to apply the composition map $g \circ f$ to an element a of A, we first apply the map f to the element a, and then we apply the map g to the resulting element $f(a)$ of B to obtain an element $g(f(a))$ of C.

The Graph of a Mapping

Let A and B be sets. To every map $f : A \to B$ from A to B there corresponds a subset $G(f)$ of the Cartesian product $A \times B$, where

$$G(f) = \{(a, b) \in A \times B : b = f(a)\}.$$

Mathematicians often refer to the subset of $A \times B$ corresponding to a map $f : A \to B$ as the graph of the map.

Example

Let $q : \mathbb{R} \to \mathbb{R}$ be the map from the set $\mathbb{R}$ of real numbers to itself defined such that $q(x) = x^2$ for all real numbers x. The graph of this map is the subset of $\mathbb{R} \times \mathbb{R}$ given by

$$\{(x, y) \in \mathbb{R} \times \mathbb{R} : y = x^2\}.$$

Note that this subset consists of the Cartesian coordinates of the points of the plane that lie on the curve that represents the graph of the given map.

Definition (Inverse of a map)

Let A and B be sets, and let $f : A \to B$ be a map from A to B. A map $g : B \to A$ from B to A is said to be the inverse of the map f if $g(f(a)) = a$ for all elements a of A and $f(g(b)) = b$ for all elements b of B.

If there exists a map $g : B \to A$ that is the inverse of $f : A \to B$, then the map f is said to be invertible and the inverse of a map $f : A \to B$ is denoted by $f^{-1} : B \to A$.

Remark

Many maps are not invertible.

The following example illustrates some of the reasons why certain maps may not be invertible.

Example

Let W be the set of all English words occurring as headwords in some specified dictionary. Let $\mathbb{N}$ denote the set of natural numbers and let $\lambda : W \to \mathbb{N}$ denote the map that sends each word to its length.
(Thus, for example, $\lambda(to) = 2$ and λ (independable) $= 12$).
This map $\lambda : W \to \mathbb{N}$ is not invertible.
For example,

$$\lambda(to) = \lambda(by) = \lambda(at) = 2$$

$$\lambda(\text{physical}) = \lambda(\text{computer}) = 8$$

i.e., there are natural numbers that are the image of more than one word.
If one were to seek to define map
$\mu : \mathbb{N} \to W$ that was the inverse of
$\lambda : W \to \mathbb{N}$ then we would have
$\mu(\lambda(\text{physical})) = \text{physical}$ and
$\mu(\lambda(\text{computer})) = \text{computer}$.
But $\mu(\lambda(\text{physical})) = \mu(8)$, and

$\mu(\lambda(\text{computer})) = \mu(8)$, and therefore the inverse map
$\mu : \mathbb{N} \to W$ would also have to satisfy
$\mu(\lambda(\text{physical})) = \mu(\lambda(\text{computer}))$, and therefore the words "physical" and "computer" would have to be identical, which is clearly not the case.
This demonstrates the impossibility of finding an inverse map to λ.

Definitions:(Injective, Surjective and Bijective Mappings)

Let A and B be sets, and let $f : A \to B$ be a map from A to B. We say that the map f is injective if $x \neq y$ for all x and y of A then $f(x) \neq f(y)$. We say that the map f is surjective if, given any element b of B, there exists some element a of A such that $f(a) = b$ or $f(A) = B$.

We say that the function f is bijective if it both injective and surjective. Thus a map is injective if and only if distinct elements of its domain get mapped to distinct elements of its codomain. A map is surjective if every element of the codomain is the image of some element of the domain.

Example.

Let $\mathbb{R}^+$ denote the set of non-negative real numbers, and let $q : \mathbb{R}^+ \to \mathbb{R}^+$ be the map given by $q(x) = x^2$ for all non-negative real numbers x. Let x and y be non-negative real numbers. If $x < y$ then $x^2 < y^2$. If $x > y$ then $x^2 > y^2$. But if $x \neq y$ then either $x < y$ or $x > y$. It follows that if $x \neq y$ then $x^2 \neq y^2$. The map $q : \mathbb{R}^+ \to \mathbb{R}^+$ is therefore injective. Also, given any non-negative real number x, there exists a non-negative real number $\sqrt{x}$ whose square is equal to x. It follows that the map $q : \mathbb{R}^+ \to \mathbb{R}^+$ is both injective and surjective. It is therefore bijective. This map also has an inverse $q^{-1} : \mathbb{R}^+ \to \mathbb{R}^+$, where $q^{-1}(x) = \sqrt{x}$ for all non-negative real numbers x.

Example

Let $s : \mathbb{R} \to \mathbb{R}$ be the map given by $s(x) = x^2$ for all real numbers x. This map is not injective. For example, -2 and 2 are distinct elements of $\mathbb{R}$, but $s(-2) = 4 = s(2)$.

Moreover, the map is not surjective, since any negative real number such as -4 is not in the range of the function. This function $s : \mathbb{R} \to \mathbb{R}$ is neither injective nor surjective.

Moreover one can easily satisfy oneself that it does not have an inverse.

(Such an inverse, were it to exist, would have to be defined for all real numbers, not merely the non-negative ones).

The next result shows that a map between sets is invertible if and only if it is a bijection.

Theorem(Invertibility of a Map)

A map $f : A \to B$ is invertible if and only if it is both injective and surjective.

Example

Let $f : [-3, 1] \to [0, 9]$ be the map defined by $f(x) = x^2$ for all $x \in [-3, 1]$, where $[-3, 1] = \{x \in \mathbb{R} : -3 \leq x \leq 1\}$ and $[0, 9] = \{x \in \mathbb{R} : 0 \leq x \leq 9\}$. The map $f : [-3, 1] \to [0, 9]$ is surjective, since for each real number y satisfying $0 \leq y \leq 9$, there exists at least one real number x satisfying $-3 \leq x \leq 1$ such that $f(x) = y$; one such real number x is given by $x = -\sqrt{y}$, where $\sqrt{y}$ denotes the positive square root of y. However, the map f is not injective. Indeed $f(1) = f(-1) = 1$. The map $f : [-3, 1] \to [0, 9]$ is therefore not bijective, and hence is not invertible.

Example

Let $f : [0, 2] \to [0, 2]$ and $g : [0, 2] \to [0, 2]$ be the map defined by

$$f(x) = \begin{cases} x^2 & \text{if } 0 \leq x \leq 1; \\ 3 - x & \text{if } 1 < x \leq 2; \end{cases}$$

$$g(x) = \begin{cases} x^2 & \text{if } 0 \leq x < 1; \\ 3 - x & \text{if } 1 \leq x \leq 2 \end{cases}$$

The map $f : [0, 2] \to [0, 2]$ is not injective since $f(1) = f(2) = 1$. This map is not surjective since there is no element x of the domain $[0, 2]$ for which $f(x) = 2$. The map f is thus not bijective, and hence is not invertible. The

map $g : [0, 2] \to [0, 2]$, on the other hand, is invertible, with inverse given by

$$g^{-1}(x) = \begin{cases} \sqrt{x} & \text{if } 0 \leq x < 1; \\ 3 - x & \text{if } 1 \leq x \leq 2. \end{cases}$$

It follows from this that the map $g : [0, 2] \to [0, 2]$ must be both injective and surjective.

Summary
We defined and gave some examples and results on mapping, its domain, its codomain, its range, its inverse, its graph and its compositions.
Special mappings are also studied such as:
(i) injective maps
(ii) surjective maps, and
(iii) bijective maps.

Post-Test
See Pre-Test at the beginning of the lecture.

Supplementary Reading

1. A.,O. Kuku, Abstract Algebra, Ibadan University Press, 1980, pp. 25-35.
2. S.A. Ilori and O. Akinyele, Elementary Abstract and Linear Algebra, Ibadan University Press, 1986. pp. 142-149, 152-156.
3. S. Lipschutz, Set Theory and Related Topics, Schaum's Outline Series.
4. F. Ayres, Modern Algebra, Schaum's Outline Series.
5. S. Lipschutz, General Topology, Schaum's Outline Series Chapter 2.
6. G. Birkoff and S. MaClane, A survey of Modern Algebra, Macmillan Co. 1965.
7. J.B. Fraleigh, A First Course in Abstract Algebra, Addison Wesley 1968.

LECTURE 4

Matrix and System of Linear Equations

Introduction

We define matrices with their notations and the basic operations of matrices. Also, we state some important properties about the inverse of a matrix. Finally, we discuss how to solve systems of linear equations using Gaussian and Gauss-Jordan Elimination.

Objective
At the end of this lecture, you should be able to
- identify matrices,
- define and manipulate with the basic operations of matrices,
- know the properties and uniqueness of inverse of a matrix, and
- solve systems of linear equations through Gaussian and Gauss-Jordan Elimination

Pre-Test

1. Let
$$A = \begin{pmatrix} 2 & 2 & -1 & -1 \\ -1 & 0 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{pmatrix}$$
Compute $I_4 - A$, $(I_4 - A)^2$, $(I_4 - A)^3$ and $(I_4 - A)^4$.

2. Let
$$A = \begin{pmatrix} 1 & t \\ \frac{1}{t} & 1 \end{pmatrix}$$

Find A^2 and A^3.

4. Compute the matrix product

$$\begin{bmatrix} x & y & 1 \end{bmatrix} \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Hence express in matrix notation the equations

(a) $x^2 + 9xy + y^2 + 8x + 5y + 2 = 0$;

(b) $\dfrac{x^2}{\alpha^2} + \dfrac{y^2}{\beta^2} = 1$;

(c) $xy = \alpha^2$;

(d) $y^2 = 4\alpha x$.

4. Reduce to row-echelon form the augmented matrix of the system of equations

$$\begin{array}{rcrcrcr} x & + & 2y & & & + & 3t = 1 \\ x & + & 2y & + & 3z & + & 3t = 3 \\ x & & & + & z & + & t = 3 \\ x & + & y & + & z & + & 2t = 1 \end{array}$$

Deduce that the system has no solution.

5. For what value of λ does the system of equations

$$\begin{array}{rcrcrcr} x & + & y & & & + & t = 4 \\ 2x & & & & & - & 4t = 7 \\ x & + & y & + & z & & = 5 \\ x & - & 3y & - & z & - & 10t = \lambda \end{array}$$

have a solution? Find the general solution when λ takes this value.

6. What conditions must the integers a, b, c satisfy in order that the system of equations

$$\begin{array}{rcrcrcr} 2w & - & x + y & - & 3z & = & a \\ w & + & x - y & & & = & b \\ 4w & + & x - y & - & 3z & = & c \end{array}$$

has integer solutions?

7. Show that the equations

$$\begin{array}{rcr} x-y & & -u-5t=\alpha \\ 2x+y & - & z-4u+t=\beta \\ x+y & + & z-4u-6t=\gamma \\ x+4y & + & 2z-8u-5t=\delta \end{array}$$

have a solution if and only if

$$8\alpha-\beta-11\gamma+5\delta=0.$$

Find the general solution when $\alpha=\beta=-1$, $\gamma=3$, $\delta=8$.

Matrices and Algebraic Operations of Matrices

Definition (Matrix)

An $m\times n$ matrix A is a rectangular array of numbers with m rows and n columns and written as

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

a_{ij} is called the (i,j)th element of A.
Note that a special case of matrix is a vector where either $m=1$ or $n=1$. If $m=1$ and $n>1$, then it is called a row vector. If $m>1$ and $n=1$, then it is called a column vector.

Now, we define the basic operations of matrices.

Definition

Let A and B be $m\times n$ matrices.

1. (equality) $A=B$ if $a_{ij}=b_{ij}$
2. (addition) $C=A+B$ if $c_{ij}=a_{ij}+b_{ij}$ and C is an $m\times n$ matrix.

3. (Scalar multiplication). Given $k \in \mathbb{R}$, $C = kA$ if $c_{ij} = ka_{ij}$ where C is an $m \times n$ matrix.

4. (product) Let C be an $n \times l$ matrix. $D = AC$ if $d_{ij} = \sum_{k=1}^{n} a_{ik}c_{kj}$ and D is an $m \times l$ matrix.

5. (transpose) $C = A^T$ if $c_{ij} = a_{ji}$ and C is an $n \times m$ matrix.

Practice Exercise

Calculate $A + 2B^T$, AB and BA using the following matrices.

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 1 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} -2 & 5 \\ 4 & -3 \\ 2 & 1 \end{bmatrix}.$$

The basic algebraic operations for matrices are as follows:

Theorem(Basic Algebraic Operations for Matrices)

Let A, B, C be matrices of appropriate sizes.

1. Addition:
   (a) $A + B = B + A$ and $A + (B + C) = (A + B) + C$
   (b) There exists a unique C such that $A + C = A$ and we denote $C = 0$.
   (c) There exists a unique C such that $A + C = 0$ and we denote $C = -A$
2. Multiplication:
   (a) $k(lA) = (kl)A$, $k(A + B) = kA + kB$, $(k + l)A = kA + lA$, and $A(kB) = k(AB) = (kA)B$ for any $k, l \in \mathbb{R}$.
   (b) $A(BC) = (AB)C$.
   (c) $(A + B)C = AC + BC$ and $C(A + B) = CA + CB$.
3. Transpose:

(a) $(A^T)^T = A$

(b) $(A + B)^T = A^T + B^T$

(c) $(AB)^T = B^T A^T$

(d) $(kA)^T = kA^T$

Practice Exercise

Calculate $(ABC)^T$ using the following matrices:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -2 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 2 & 2 \\ 3 & -1 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 2 \\ 3 & -1 \end{bmatrix}[,$$

There are some important special types of matrices.

Definitions (square, symmetric, diagonal and identity matrices)

Let A be an $m \times n$ matrix.

1. A is called a square matrix if $n = m$.
2. A is called symmetric if $A^T = A$.
3. A square matrix A is called a diagonal matrix if $a_{ij} = 0$ for $i \neq j$. A is called upper triangular if $a_{ij} = 0$ for $i > j$ and called lower triangular if $a_{ij} = 0$ for $i < j$.
4. A diagonal matrix A is called an identity matrix if $a_{ij} = 1$ for $i = j$ and is denoted by I_n.

In particular, we have $AI_n = I_n A = A$ for any square matrix A. For a square matrix, there is another operator called trace.

Definition (Trace of a matrix)

If A is an $n \times n$ matrix, then the trace of A denoted by $tr(A)$ is defined as the sum of all the main diagonal elements of A. That is,

$$tr(A) = \sum_{i=1}^{n} a_{ii}$$

Some useful facts about trace operators are given below.

Theorem(Trace Operators)
Let A and B be matrices of appropriate sizes.

1. $tr(kA) = ktr(A)$ for any $k \in \mathbb{R}$
2. $tr(A + B) = tr(A) + tr(B)$
3. $tr(AB) = tr(BA)$.
4. $tr(A^T) = tr(A)$.
5. $tr(A^T A) \geq 0$.

If we start out with an $m \times n$ matrix and delete some but not all, of its rows or columns, then we obtain a submatrix. For example, $C = \begin{bmatrix} 3 & 2 \\ 3 & -1 \end{bmatrix}$ is a submatrix of the following matrix.

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

A matrix can be partitioned into submatrices, and such a matrix is called partitioned matrices. Partitioned matrices can be manipulated in the same way (called block manipulation) provided that submatrices are of appropriate sizes.

Example

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}.$$

Definition(Non Singular Matrix)

An $n \times n$ matrix A is called nonsingular or invertible if there exists an $n \times n$ matrix B such that $AB = BA = I_n$. We call such B an inverse of A. Otherwise, A is called singular or noninvertible.

Practice Exercise

Show that B is an inverse of A (or A is an inverse of B).

$$A = \begin{bmatrix} 2 & 3 \\ 2 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} -1 & \frac{3}{2} \\ 1 & -1 \end{bmatrix}.$$

We state without proof some important properties about the inverse of a matrix.

Theorem(Uniqueness of Inverse of a Matrix)

The inverse of a matrix, if it exists, is unique. We denote the unique inverse of A by A^{-1}.

Theorem(Inverse of Product of Matrices)

Let A and B be nonsingular $n \times n$ matrices.

1. AB is nonsingular and $(AB)^{-1} = B^{-1}A^{-1}$.
2. A^{-1} is nonsingular and $(A^{-1})^{-1} = A$.
3. $(A^T)^{-1} = (A^{-1})^T$.

System of Linear Equations

One application of inverting a matrix is to solve a system of linear equations. In fact, matrices can be motivated in terms of linear equations. Consider a set of m linear equations of the form

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ &\vdots \qquad\qquad \vdots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{aligned}$$

Then, its matrix representation is $Y = AX$ where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

We call A a coefficient matrix.
With this notation, we can see that A^{-1} (provided that A is nonsingular) solves this system since we obtain $X = A^{-1}Y$ by premultiplying the equation by A^{-1}.

Practice Exercise

Confirm that

$$\begin{bmatrix} 2 & 3 \\ 2 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & \frac{3}{2} \\ 1 & -1 \end{bmatrix}$$

Solve the following system of linear equations by using the inverse of matrix.

$$\begin{aligned} 2x_1 + 3x_2 &= 1 \\ 2x_1 + 2x_2 &= 2 \end{aligned}$$

Since we do not yet know how to find the inverse of a matrix in general, we rely on "elementary row (column) operations on matrices" to solve a system of linear equations.

We now introduce the following definitions.

Definition (Elementary row(column) operations on a matrix)

Elementary row (column) operations on an $m \times n$ matrix A includes the following:

1. Interchange rows (columns) r and s of A
2. Multiply row(column) r of A by a nonzero scalar $k \neq 0$.
3. Add k times row (column) r of A to row (column) s of A where $r \neq s$.

An $m \times n$ matrix A is said to be row(column) equivalent to $m \times n$ matrix B if B can be obtained by applying a finite sequence of elementary row (column) operations to A.

Practice Exercise

Show that A is row equivalent to B.

$$A = \begin{bmatrix} 1 & 2 & 4 & 3 \\ 2 & 1 & 3 & 2 \\ 1 & -1 & 2 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 4 & 8 & 6 \\ 1 & -1 & 2 & 3 \\ 4 & -1 & 7 & 8 \end{bmatrix}.$$

Now, we can use these operations to characterise systems of linear equations.

Theorem(Characterisation System of Linear Equations)

Let $AX = B$ and $CX = D$ be two linear systems with m equations and n unknowns. If the augmented matrices $[AB]$ and $[CD]$ are row equivalent, then the linear systems have the same solutions.

Finally, to solve systems of linear equations using elementary row(column) operations, we need one more concept.

Definition (Reduced row echelon form)

An $m \times n$ matrix A is said to be in reduced row echelon form if it satisfies the following properties.

1. All rows consisting entirely of zeros, if any, are at the bottom of the matrix.

2. By reading from left to right, the first nonzero entry in each row that does not consist entirely of zeros is a 1, called the leading entry of its row.

3. If rows i and $i+1$ are two successive rows that do not consist entirely of zeros, then the leading entry of row $i+1$ is to the right of the leading entry of row i.

4. If a column contains a leading entry of some row, then all other entries in that column are zero.

If A satisfies 1,2 and 3, but not 4, then it is said to be in row echelon form. A similar definition can be applied to (reduced) column echelon form.

Example

A is in row echelon form, and B is in reduced row echelon form

$$A = \begin{bmatrix} 1 & 5 & 0 & 2 & -2 & 4 \\ 0 & 1 & 0 & 3 & 4 & 8 \\ 0 & 0 & 0 & 1 & 7 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 & 0 & 0 & -2 & 4 \\ 0 & 1 & 0 & 0 & 4 & 8 \\ 0 & 0 & 0 & 1 & 7 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Finally, we define two methods that can be used to solve systems of linear equations.

Theorem(Solve System of Linear Equations by Gaussian Elimination)

A system of linear equations $AX = Y$ can be solved by using Gaussian (Gauss-Jordan) elimination, which consists of the following two steps:

1. Use elementary operations to transform the augmented matrix $[AB]$ to the matric $[CD]$ in (reduced) row echelon form.

2. Solve the linear system corresponding to the augmented matrix $[CD]$ using back substituting.

Practice Exercise

Solve the following system of linear equations using the Gaussian elimination

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ 2x_1 - x_2 + x_3 &= 8 \\ 3x_1 - x_3 &= 3 \end{aligned}$$

Summary

We discussed the basic laws of matrix algebra and introduced Gaussian (Gauss-Jordan) elimination to solve systems of linear equations.

Post-Test

See Pre-Test at the beginning of this lecture.

Supplementary Reading

1. M. Artin, Algebra: Matrix Operations (Chpter 1). Prentice-Hall, Upper Saddle River, New Jersey 07458, 1991.

2. H.A. Elliot, K.D. Fryer, J.C. Gardner and N.J. Hill, Vectors and Matrices. (Chapters 6 and 8).
Holt, Rinehart and Winston of Canada Limited, 1966.

3. S.A. Ilori and O. Akinyele, Elementary Abstract and Linear Algebra, Ibadan University Press, 1986. pp. 372-411.

4. E.H. Connell, Elements of Abstract and Linear Algebra.
(http://www.math.miami.edu/∼ec/book/)

LECTURE 5

Determinant and Inverse of Matrix

Introduction

In this lecture, we will define and examine some basic properties of determinant. Also, we shall show how to use cofactors to calculate the inverse of a matrix.

Finally, we shall introduce Cramer's Rule to solve a system of linear equations.

Objective
At the end of this lecture you should be able to:
- compute the determinants of matrices,
- apply cofactor to calculate the inverse of a matrix, and
- use Cramer's Rule to solve a system of linear equations.

Pre-Test

1. Find the inverses of the following matrices:

$$\begin{bmatrix} 3 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}; \quad \begin{bmatrix} 5 & 3 & 2 \\ 2 & 3 & 1 \\ 7 & 5 & 3 \end{bmatrix}; \quad \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 3 \end{bmatrix}.$$

2. Determine the values of the real number x for which the matrix

$$\begin{bmatrix} x & 2 & 0 & 3 \\ 1 & 2 & 3 & 3 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 3 \end{bmatrix}$$

is invertible.

3. Given the matrices

$$A = \begin{bmatrix} b+8c & 2c-2b & 4b-4c \\ 4c-4a & c+8b & 2a-2c \\ 2b-2a & 4a-4b & a+8b \end{bmatrix}, \quad P = \begin{bmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix},$$

find P^{-1} and evaluate $P^{-1}AP$. Hence evaluate $\det A$.

4. If A and B are square matrices of the same order prove that

$$\det \begin{bmatrix} A & B \\ B & A \end{bmatrix} = \det(A+B)\det(A-B).$$

5. If

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ a & x & b & c \\ a^2 & x^2 & b^2 & c^2 \\ a^3 & x^3 & b^3 & c^3 \end{bmatrix}$$

express $\det A$ as a product of factors.

6. Answer the following true or false and explain your reasoning.

   (a) If A and B are $n \times n$ matrices and $AB = 0$, then either $A = 0$ or $B = 0$

   (b) If A is a matrix with $A^2 = I_n$ then either $\det(A) = 1$ or $\det(A) = -1$.

Definition (Permutation)

A permutation of a finite set of integers $S = \{1, 2, \ldots, n\}$ is a bijective map $f : S \to S$. A permutation is said to have an inversion if a larger integer proceeds a smaller one. A permutation is called even (odd) if the total number of inversions is even (odd).

That is, if $S = \{1, 2, 3\}$, then f defined by $f(1) = 3$, $f(2) = 2$, $f(3) = 1$ is an odd permutation.

Now, we are ready to define determinant of a matrix.

Definition (Determinant)

Let A be an $n \times n$ matrix. Then, the determinant of A denoted by $|A|$ or $\det(A)$ is

$$\sum(\pm)a_{1f(1)}a_{2f(2)}\dots a_{nf(n)}$$

where the summation is over all permutations $f : S \to S$ with $S = \{1, 2, \dots, n\}$. The sign is $+(-)$ if the corresponding permutation is even (odd).

Now, we compute the determinants of the following matrices. It should be noted that there is no easy method for computing determinants for $n > 3$.

Practice Exercise

What are the determinants of 1×1, 2×2 and 3×3 matrices?

We examine some basic properties of determinants. In particular, there is an important relationship between the singularity and the determinant of a matrix.

Theorem(Relationship between the Singularity and the Determinant of a Matrix)

Let A and B be $n \times n$ matrices.

1. $|I_n| = 1$ and $|-I_n| = (-1)^n$.
2. $|kA| = k^n|A|$ for $k \in \mathbb{R}$.
3. $|A| = |A^T|$.
4. A is nonsingular if and only if $|A| \neq 0$
5. $|AB| = |A||B|$.
6. If A is nonsingular, then $|A^{-1}| = |A|^{-1}$.

According to Definition of determinant, computing the determinant of an $n \times n$ matrix can be very cumbersome if n is large. We now develop method which reduces the problem to the computation of the determinant of an $(n-1) \times (n-1)$ matrix so that we can repeat the process until we get to a 2×2 matrix.

Definition (Minor and Cofactor of a Matrix)

Let A be an $n \times n$ matrix.

1. Let M_{ij} be the $(n-1) \times (n-1)$ submatrix of A obtained by deleting the ith row and jth column of A. Then, $|M_{ij}|$ is called the minor of a_{ij}.
2. The cofactor of A_{ij} of a_{ij} is defined as

$$A_{ij} = (-1)^{i+j}|M_{ij}|$$

Now, the following theorem gives us a new method for computing determinants.

Theorem (Cofactor Expansion)

Let A be an $n \times n$ matrix. Then, for any i and j, $|A| = \sum_{j=1}^{n} a_{ij}A_{ij}$ and $|A| = \sum_{i=1}^{n} a_{ij}A_{ij}$.

Practice Exercise

Find the determinant of the following matrix using cofactor expansion.

$$A = \begin{bmatrix} 1 & 2 & -3 & 4 \\ -4 & 2 & 1 & 3 \\ 3 & 0 & 0 & -3 \\ 2 & 0 & -2 & 3 \end{bmatrix}$$

The following will show how one can use cofactors to calculate the inverse of a matrix.

Definition (Adjoint of a matrix)

Let A be an $n \times n$ matrix. The adjoint of A, $adj\ A$, is the matrix whose (i, j) element is the cofactor A_{ji} of $_{ji}$. That is,

$$adj\ A = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}.$$

Practice Exercise

Compute the adjoint of the following matrix

$$A = \begin{bmatrix} 3 & -2 & 1 \\ 5 & 6 & 2 \\ 1 & 0 & -3 \end{bmatrix}.$$

Inverse of a Matrix

Finally, the inverse of a square matrix can be written as follows.

Theorem(Inverse of a Matrix)

If A is an $n \times n$ matrix and $|A| \neq 0$, then

$$A^{-1} = \frac{1}{|A|} adj A.$$

The theorem illustrates why $|A| \neq 0$ is required for A^{-1} to exist.

Practice Exercise

Compute the inverse of

$$A = \begin{bmatrix} 3 & -2 & 1 \\ 5 & 6 & 2 \\ 1 & 0 & -3 \end{bmatrix}.$$

Now, you can solve a system of linear equations provided its solution exists i.e., the inverse of the coefficient matrix exists.

Cramer's Rule

Theorem

Consider a system of n linear equations in n unknown parameters with the coefficient matrix A so that we can write $Y = AX$

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots &\quad\quad \vdots \quad\quad \vdots \\ y_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{aligned}$$

If $|A| \neq 0$, then the system has the unique solution.

$$x_1 = \frac{|A_1|}{|A|}, x_2 = \frac{|A_2|}{|A|}, \cdots, x_n = \frac{|A_n|}{|A|},$$

where A_i is the matrix obtained from A by replacing its ith column by Y.

Practice Exercise

Apply the Cramer's Rule to the following system of linear equations

$$\begin{aligned} -2x_1 + 3x_2 - x_3 &= 1 \\ x_1 + 2x_2 - x_3 &= 4 \\ -2x_1 - x_2 + x_3 &= -3 \end{aligned}$$

Summary

Properties of determinant are studied.
Next, we used cofactor expansion to determine determinant as well as coefactors to calculate the inverse of a matrix. Finally, Cramer's Rule are stated with applications.

Post-Test

See Pre-Test at the beginning of the lecture.

Supplementary Reading

1. M. Artin, Algebra: Matrix Operations (Chpter 1). Prentice-Hall, Upper Saddle River, New Jersey 07458, 1991.

2. H.A. Elliot, K.D. Fryer, J.C. Gardner and N.J. Hill, Vectors and Matrices. (Chapters 6 and 8).
Holt, Rinehart and Winston of Canada Limited, 1966.

3. S.A. Ilori and O. Akinyele, Elementary Abstract and Linear Algebra, Ibadan University Press, 1986. pp. 231-258, 300-323.

4. E.H. Connell, Elements of Abstract and Linear Algebra.
(http://www.math.miami.edu/∼ec/book/)

LECTURE 6

Vector Spaces

Introduction

We shall consider vector spaces over a field ($\mathbb{R}$ $\mathbb{Q}$ or $\mathbb{C}$) and important subsets of a vector space called subspaces.

Objectives
At the end of this lecture you should be able to:
- define a vector space and its subspaces,
- show whether a structure is a vector space or not, and
- show whether a given non-empty subset of a vector space is a subspace

Pre-Test

1. Determine whether or not the following subsets of $\mathbb{R}^4$ are subspaces:
   (a) $U = \{(a, b, c, d) : a + b = 1\}$
   (b) $U = \{(a, b, c, d) : a^2 + b^2 = 0\}$.
2. Prove that $V = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$ is a vector space over $\mathbb{Q}$.
3. Let V be the real vector space of continuous functions $f : \mathbb{R} \to \mathbb{R}$. Which of the following subsets are subspaces of V?
   (a) $W_1 = \{f \in V : f(\frac{1}{2}) \in \mathbb{Q}\}$;
   (b) $W_2 = \{f \in V : f(\frac{1}{2}) = f(1)\}$;
   (c) $W_3 = \{f \in V : f(\frac{1}{2}) = 0\}$;
   (d) $W_4 = \{f \in V : Df(\frac{1}{2}) = 1\}$.

Find $W_i \cap W_j$ in the cases where W_i and W_j are subspaces.

4. Which of the following statements are true? Give a proof for those which are true and a counter example for those which are false.

   (a) $\{(x_1, x_2) \in \mathbb{R}^2 : x_1 < x_2\}$ is a subspace of $\mathbb{R}^2$.

   (b) The subspace of $\mathbb{R}^3$ spanned by $\{(1,2,1),(2,2,1)\}$ is $\{(x+y, 2y, y) : x, y \in \mathbb{R}\}$.

   (c) The subspace of $\mathbb{R}^3$ spanned by $\{(1,2,1),(2,2,1)\}$ is $\{(2x, 2x+2y, x+y) : x, y \in \mathbb{R}\}$.

5. Let U be the subspace of $\mathbb{R}^4$ spanned by

$$X = \{(2,2,1,3),(7,5,5,5),(3,2,2,1),(2,1,2,1)\}.$$

Given that $x = (6+\lambda, 1+\lambda, -1+\lambda, 2+\lambda)$ belongs to U, find λ. For this value of λ, does x have a unique expression as a linear combination of the vectors of X?

Definition(A Vector Space)

A vector space V is a set with two operations $+$ and $\cdot$ that satisfy the following properties.

(a) If u and v are elements of V, then $u+v$ is an element of V (closure under $+$)

1. $u + v = v + u$
2. $u + (v + w) = (u + v) + w$
3. There is an element o in V such that

$$u + o = o + u = u$$

4. For every u in V there is an element $-u$ with

$$u + (-u) = 0$$

(b) If u is in V and c is a real number then $c \cdot u$ is in V (closure under $\cdot$)

1. $c \cdot (u + v) = c \cdot u + c \cdot v$
2. $(c + d) \cdot u = c \cdot u + d \cdot u$
3. $c \cdot (d \cdot u) = (cd) \cdot u$
4. $1 \cdot u = u$

Practice Exercise

Verify the properties of vector space on the set

$$V = \mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n-\text{times}} = \{(x_1, x_2, \ldots, , x_n) : x_i \in \mathbb{R}, \quad i = 1, \ldots, n\}$$

Remark

You will recognize these properties of vectors in $\mathbb{R}^n$, however there is a large class of vector spaces that do not look at all like $\mathbb{R}^n$.

Examples of Vector Spaces

Example 1 P_2

Consider the set P_2 of polynomials of degree less than or equal to 2. Define $+$ to be polynomial addition:
For $(a_1t^2 + b_1t + c_1) \in P_2$ and $(a_2t^2 + b_2t + c_2) \in P_2$ we have

$$\begin{aligned}&(a_1t^2 + b_1t + c_1) + (a_2t^2 + b_2t + c_2)\\ &= (a_1 + a_2)t^2 + (b_1 + b_2)t + (c_1 + c_2)\end{aligned}$$

for all a_1, b_1, c_1, a_2, b_2 and c_2 in $\mathbb{R}$ and $\cdot$ is defined by

$$k \cdot (at^2 + bt + c) = (ka)t^2 + (kb)t + (kc)$$

for all k, a, b and c in $\mathbb{R}$.
This is a vector space. Most of the properties clearly hold. We will demonstrate a few of the properties. For examle the 0 vector is the zero polynomial (0) i.e., $0t^2 + 0t + 0 \in (0)$. We have

$$(at^2 + bt + c) + 0 = 0 + (at^2 + bt + c) = at^2 + bt + c$$

Property $b2$ holds since for all r, s, a, b and c in $\mathbb{R}$,

$$\begin{aligned}(r+s)\cdot(at^2+bt+c) &= (r+s)at^2+(r+s)bt+(r+s)c\\ &= (ra+sa)t^2+(rb+sb)t+(rc+sc)\\ &= (ra)t^2+(sa)t^2+(rb)t+(sb)t+rc+sc\\ &= (ra)t^2+(rb)t+rc+(sa)t^2+(sb)t+sc\\ &= r(at^2+bt+c)+s(at^2+bt+c)\end{aligned}$$

Practice Exercise

We can generalize Example 1 and let P_n be the set of all polynomials of degree less than n. We define $+$ to mean polynomial addition and $\cdot$ to be scalar multiplication as in Example 1. This is a vector space as you can check.

Example 2 $M_{2\times 3}$

Consider the set $M_{2\times 3}$ of 2×3 matrices and let $+$ be defined by matrix addition and $\cdot$ be defined by matrix scalar multiplication. Then $M_{2\times 3}$ is a vector space. We have stated all of the required properties previously.

Practice Exercise

We can generalize Example 2 by letting $M_{m\times n}$ be the set of all $m\times n$ matrices with matrix addition and scalar multiplication as before.

Example 3

Consider the set V of all differentiable functions f such that $f'(1)=0$. Let $+$ be defined as addition of functions and $\cdot$ be defined as regular scalar multiplication. This is a vector space. We will demonstrate a few of the properties. Let f and g be elements of this set. Then

$$f'(1)=g'(1)=0.$$

To show additive closure, we have

$$(f+g)'(1)=f'(1)+g'(1)=0+0=0$$

so that $f+g$ is in V.
To show multiplicative closure we have

$$(cf)'(1)=c(f'(1))=c(0)=0.$$

The rest of the properties follow from the properties of function arithmetic and derivatives.

Example 4

Let S be the set of ordered pairs in $\mathbb{R}^2$ with $+$ defined by

$$(x_1, y_1) + (x_2, y_2) = (x_1 + 2x_2, y_1 + 2y_2)$$

and $\cdot$ defined by

$$c(x, y) = (cx, cy)$$

then S is not a vector space, since property $a1$ fails. For example, $(2, 3) + (4, 5) = (10, 13)$ but $(4, 5) + (2, 3) = (8, 11)$.

Definition (Subspace)

Let V be a vector space and let S be a subset of V such that S is a vector space with the same $+$ and $\cdot$ from V. Then S is called a subspace of V.

Remark

Every vector space V contains at least two subspaces, namely V and the set $\{0\}$.

Example

Let V be the vector space $\mathbb{R}^3$ and let S be the set of points that lie on the plane

$$z = x - y$$

Then S is a subspace of V.
This is true since S is closed under $+$ and $\cdot$.
A point belonging to S has the form

$$(x, y, x - y)$$

If $(x_1, y_1, x_1 - y_1)$ and $(x_2, y_2, x_2 - y_2)$ are in S then

$$\begin{aligned}(x_1, y_1, x_1 - y_1) + (x_2, y_2, x_2 - y_2) &= (x_1 + x_2, y_1 + y_2, x_1 - y_1 + x_2 - y_2) \\ &= (x_1 + x_2, y_1 + y_2, (x_1 + x_2) - (y_1 + y_2))\end{aligned}$$

is in S. We also have

$$c(x_1, y_1, x_1 - y_1) = (cx_1, cy_1, cx_1 - cy_1)$$

is in S.
The rest of the properties follow immediately since they are true in V. In fact, the two closure properties are all we need to show when we want to check that any subset S is a subspace of any vector space V.
Theorem(Subset of a Vector Space)

Let V be a vector space and S be a subset of V. If S is closed under $+$ and $\cdot$ then S is a subspace of V.
The proof of this theorem only involves noticing that the properties are all true in S by virtue of being true in V.

Example

Set V be the vector space of all differentiable functions and let S be the subset of V such that for any f in V.

$$f'(0) = 1.$$

Then S is not a subspace of V since if f and g are in S, then

$$(f + g)'(0) = f'(0) + g'(0) = 1 + 1 = 2$$

Hence S is not closed under $+$.

Example

Let S be the subset of $M_{2\times 2}$ of trace 0, that is the sum of the diagonal entries is zero. Then S is a subspace of $M_{2\times 2}$.
Elements of S have the form

$$\begin{pmatrix} x & y \\ z & -x \end{pmatrix}$$

so if

$$A = \begin{pmatrix} x_1 & y_1 \\ z_1 & -x_1 \end{pmatrix}, \quad B = \begin{pmatrix} x_2 & y_2 \\ z_2 & -x_2 \end{pmatrix}$$

then

$$A + B = \begin{pmatrix} x_1 + x_2 & y_1 + y_2 \\ z_1 + z_2 & -x_1 - x_2 \end{pmatrix}$$

has zero trace. And

$$cA = c \begin{pmatrix} x_1 & y_1 \\ z_1 & -x_1 \end{pmatrix} = \begin{pmatrix} cx_1 & cy_1 \\ cz_1 & -cx_1 \end{pmatrix}$$

also has trace zero. Hence S is closed under $+$ and $\cdot$.
We can conclude that S is a subspace of V.

Example (The Range and Null space of a Matrix)

We define subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$ as follows.
Let A be an m and n matrix.
Define the null space of A to be the subspace considering of vectors x in $\mathbb{R}^n$ with the property.

$$Ax = 0$$

and define the range of A to be the subspace consisting of vectors y in $\mathbb{R}^m$ such that there is a b in $\mathbb{R}^n$ with

$$Ab = y.$$

Example (Linear Combinations and Span)

Suppose that $V = P_2$ and let $f(t) = t^2 - t$ and $g(t) = t + 1$.
Let S be the subset of P_2 that consists of all polynomials of the form

$$c_1 f + c_2 g$$

where c_1 and c_2 are constants. Then S is called the span of f and g and is a subspace of P_2.
We will show closure under $\cdot$
Closure under $+$ is also not difficult to show and is left to the reader to check.
Let $u = c_1 f + c_2 g$ and c be a constant. Then

$$\begin{aligned} cu &= cc_1 f + cc_2 g \\ &= af + bg \end{aligned}$$

so is in S.
In general, if V is a vector space and $S = \{v_1, v_2, \ldots, v_n\}$ is a subset of V, then we call

$$c_1v_1 + c_2v_2 + \cdots + c_nv_n$$

a linear combination of S. The set of all linear combinations of S is called the span of S and is a subspace of V.

Summary
We introduced vector spaces and their subspaces and gave various examples of the structures.

Post-Test
See Pre-Test at the beginning of the lecture.

Supplementary Reading

1. M. Artin, Algebra: Matrix Operations (Chpter 1). Prentice-Hall, Upper Saddle River, New Jersey 07458, 1991.

2. S.A. Ilori and O. Akinyele, Elementary Abstract and Linear Algebra, Ibadan University Press, 1986. pp. 329-418.

3. D.C. Lay, Linear Algebra and Its Applications (3rd ed.), Addision Wesley, ISBN 978-0321287137 (August 22, 2005).

4. E.H. Connell, Elements of Abstract and Linear Algebra. (http://www/math.miami.edu/∼ec/book/)

5. J. Hefferon, Linear Algebra (http://joshua,smcvt.edu/linalg.html/) excellent textbook with complete solutions manual.

LECTURE 7

Spanning Sets and Linear Independence

Introduction

We have seen in the last discussion that the span of vectors $v_1, v_2, \ldots, v_n$ is the set of linear combinations

$$c_1v_1 + c_2v_2 + \cdots + c_nv_n$$

and that this is a vector space. We now take this idea further and discuss linearly dependent sets of a vector space.

Objectives
At the end of this lecture you should be able to do the following:
• determine the spanning sets of a vector space
• determine the smaller subsets of the spanning sets that can be written as a linear combination of the others, and
• state and prove the necessary and sufficient condition for a subset of a vector space to be linearly dependence.

Pre-Test

1. If x, y, z are vectors in $\mathbb{R}^n$, prove that

$$\langle x, y, z\rangle = \langle x + y, x + z, y + z\rangle$$

2. Determine if

$$x_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix} \quad \text{and} \quad x_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 3 \end{bmatrix}$$

are linearly independent in $\mathbb{R}^4$.

3. For which real numbers λ are the following vectors linearly independent in $\mathbb{R}^3$?

$$x_1 = \begin{bmatrix} \lambda \\ -1 \\ -1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} -1 \\ \lambda \\ -1 \end{bmatrix} \quad \text{and} \quad x_3 = \begin{bmatrix} -1 \\ -1 \\ \lambda \end{bmatrix}$$

4. Let U be the subspace of $\mathbb{R}^4$ spanned by

$$S = \{(2,2,1,3), (7,5,5,5), (3,2,2,1), (2,1,2,1)\}$$

Given that $u = (6+\lambda, 1+\lambda, -1+\lambda, 2+\lambda)$ belongs to U, find λ. For this value of λ, does u have a unique expression as a linear combination of the vectors of S?

5. Let $\mathbb{Z}_3$ be the field of integers modulo 3. Consider the $\mathbb{Z}_3$-vector space $\mathbb{Z}_3^3 = \{(a,b,c) : a,b,c \in \mathbb{Z}_3\}$. Which of the following subsets are linearly independent?

   (a) $A_1 = \{(1,2,0), (2,1,0)\}$

   (b) $A_2 = \{(1,1,1), (1,0,1), (1,0,0), (0,0,1)\}$

   (c) $A_3 = \{(1,2,0), (1,1,1), (2,0,1)\}$

   (d) $A_4 = \{(1,0,1), (1,1,0), (0,1,1)\}$

6. Let V be a real vector space and suppose that $v_1, \ldots, v_k$ are linearly independent vectors in V. If $v = \sum_{i=1}^{k} a_i v_i$ where each $a_i \in \mathbb{R}$, prove that $v - v_1, \ldots, v - v_k$ are linearly independent if and only if $\sum_{i=1}^{k} a_i \neq 1$.

If V is a vector space and $S = \{v_1, v_2, \ldots, v_k\}$ is a subset of V, then is Span(S) equal to V?

Definition (Span)

Let V be a vector space and let $S = \{v_1, v_2, \ldots, v_n\}$ be a subset of V. We say that S spans V if every vector v in V can be written as a linear combination of vectors in S.

$$v = c_1v_1 + c_2v_2 + \cdots + c_nv_n.$$

Example

Show that the set

$$S = \{(0,1,1), (1,0,1), (1,1,0)\}$$

spans $\mathbb{R}^3$ and write the vector $(2,4,8)$ as a linear combination of vectors in S.

Solution

A vector in $\mathbb{R}^3$ has the form

$$v = (x, y, z)$$

Hence we need to show that every such v can be written as

$$\begin{aligned}(x,y,z) &= c_1(0,1,1) + c_2(1,0,1) + c_3(1,1,0)\\ &= (c_2 + c_2, c_1 + c_3, c_1 + c_2)\end{aligned}$$

This corresponds to the system of equations

$$\begin{array}{rcrcrcl} & & c_2 & + & c_3 & = & x \\ c_1 & + & & & c_3 & = & y \\ c_1 & + & c_2 & & & = & z \end{array}$$

which can be written in matrix form

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

We can write this as

$$Ac = b$$

Notice that

$$\det(A) = 2.$$

Hence A is nonsingular and

$$c = A^{-1}b$$

So that a nontrivial solution exists. To write $(2, 4, 8)$ as a linear combination of vectors in S, we find that

$$A^{-1} = \begin{pmatrix} -0.5 & 0.5 & 0.5 \\ 0.5 & -0.5 & 0.5 \\ 0.5 & 0.5 & -0.5 \end{pmatrix}$$

so that

$$C = \begin{pmatrix} -0.5 & 0.5 & 0.5 \\ 0.5 & -0.5 & 0.5 \\ 0.5 & 0.5 & -0.5 \end{pmatrix} \begin{pmatrix} 2 \\ 4 \\ 8 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ -1 \end{pmatrix}.$$

We have

$$(2, 4, 8) = 5(0, 1, 1) + 3(1, 0, 1) + (-1)(1, 1, 0)$$

Example

Show that if

$$v_1 = t + 2 \quad \text{and} \quad v_2 = t^2 + 1$$

and

$$S = \{v_1, v_2\}$$

then S does not span P_2.

Solution

A general element of P_2 is of the form

$$v = at^2 + bt + c.$$

We set

$$v = c_1v_1 + c_2v_2$$

or

$$\begin{aligned} at^2 + bt + c &= c_1(t + 2) + c_2(t^2 + 1) \\ &= c_2t^2 + c_1t + 2c_1 + c_2 \end{aligned}$$

Equating coefficients gives

$$\begin{aligned} a &= c_2 \\ b &= c_1 \\ c &= 2c_1 + c_2 \end{aligned}$$

Notice that if

$$a = 1, \quad b = 1, \quad c = 1$$

there is no solution to this. Hence S does not span V.

Example

Let

$$A = \begin{pmatrix} 1 & 2 & 1 & 4 \\ 0 & 2 & 8 & 10 \\ 1 & 4 & 9 & 14 \\ -1 & 0 & 7 & 6 \end{pmatrix}.$$

Find a spanning set for the null space of A.

Solution

We want the set of all vectors x with

$$Ax = 0.$$

We find that the row reduced echelon form (rref) of A is

$$\begin{pmatrix} 1 & 0 & -7 & -6 \\ 0 & 1 & 4 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The parametric equations are

$$\begin{aligned} x_1 &= 7s + 6t \\ x_2 &= -4s - 5t \\ x_3 &= s \\ x_4 &= t. \end{aligned}$$

We can get the span in the following way. We first let $s = 1$ and $t = 0$ to get

$$v_1 = (7, -4, 1, 0)$$

and let $s = 0$ and $t = 1$ to get

$$v_2 = (6, -5, 0, 1)$$

If we let $S = \{v_1, v_2\}$ then S spans the null space of A.

Linear Independence

We now know how to find out if a collection of vectors span a vector space. It should be clear that if $S = \{v_1, v_2, \ldots, v_n\}$ then Span(S) is spanned by S. The question that we next ask is, are there any redundancies. That is, is there a smaller subset of S that also span Span(S). If so, then one of the vectors can be written as a linear combination of the others.

$$v_i = c_1v_1 + c_2v_2 + \cdots + c_{i-1}v_{i-1} + c_{i+1}v_{i+1} + \cdots + c_nv_n.$$

If this is the case then we call S a linearly dependent set. Otherwise, we say that S is linearly independent. There is another way of checking that a set of vectors are linearly dependent.

Theorem

Let $S = \{v_1, v_2, \ldots, v_n\}$ be a set of vectors, then S is linearly dependent if and only if 0 is a nontrivial linear combination of vectors in S. That is, there are constants $c_1, \ldots, c_n$ with at least one of the constants nonzero with

$$c_1v_1 + c_2v_2 + \cdots + c_nv_n = 0.$$

Proof

Suppose that S is linearly dependent, then

$$v_i = c_1v_1 + c_2v_2 + \cdots + c_{i-1}v_{i-1} + c_{i-1}v_{i+1} + \cdots + c_nv_n$$

Subtracting v_i from both sides, we get

$$c_1v_1 + c_2v_2 + \cdots + c_{i-1}v_{i-1} + v_i + c_{i+1}v_{i+1} + \cdots + c_nv_n = 0.$$

In the above equation $c_i = 1$ which is nonzero, so that 0 is a nontrivial linear combination of vectors in S.
Now let

$$c_1v_1 + c_2v_2 + \cdots + c_{i-1}v_{i-1} + c_iv_i + c_{i+1}v_{i+1} + \cdots + c_nv_n = 0$$

with c_i nonzero. Divide both sides of the equation by c_i and let $a_j = -\frac{c_j}{c_i}$ to get

$$-a_1v_1 - a_2v_2 - \cdots - a_{i-1}v_{i-1} + v_i - a_{i+1}v_{i+1} - \cdots - a_nv_n = 0$$

Finally move all the terms to the other right side of the equation to get

$$v_i = a_1v_1 + a_2v_2 + \cdots + a_{i-1}v_{i-1} + a_{i+1}v_{i+1} + \cdots + a_nv_n$$

Example

Show that the set of vectors

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

are linearly independent.

Solution

We write

$$c_1(1, 1, 3, 4) + c_2(0, 2, 3, 1) + c_3(4, 0, 0, 2) = 0$$

We get four equations

$$\begin{array}{rcrcrcl} c_1 & + & 4c_3 & & & = & 0 \\ c_1 & + & 2c_2 & & & = & 0 \\ 3c_1 & + & 3c_2 & & & = & 0 \\ 4c_1 & + & c_2 & + & 2c_3 & = & 0 \end{array}$$

The matrix corresponding to this homogeneous system is

$$A = \begin{pmatrix} 1 & 0 & 4 \\ 1 & 2 & 0 \\ 3 & 3 & 0 \\ 4 & 1 & 2 \end{pmatrix}$$

and row reduced echelon form of A i.e.,

$$rref(A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Hence,

$$c_1 = c_2 = c_3 = 0$$

and we can conclude that the vectors are linearly independent.

Example

Let $S = \{\cos^2 t, \sin^2 t, 4\}$ then S is a linearly dependent set of vectors since

$$4 = 4\cos^2 t + 4\sin^2 t.$$

Summary

We have considered the concepts: span, linear combination of vectors, and linearly dependence (or independence) vectors and gave various examples to illustrate.

Post-Test

See Pre-Test at the beginning of the lecture.

Supplementary Reading

1. S.A. Ilori and O. Akinyele, Elementary Abstract and Linear Algebra, Ibadan University Press, 1986, pp. 329-371.

1. 2.] D.C. Lay, Linear Algebra and Its Applications (3rd ed.), Addision Wesley, ISBN 978-0321287137 (August 22, 2005).

3. E.H. Connell, Elements of Abstract and Linear Algebra. (http://www/math.miami.edu/∼ec/book/)

4. J. Hefferon, Linear Algebra (http://joshua,smcvt.edu/linalg.html/) excellent textbook with complete solutions manual.

LECTURE 8

Basis and Dimension

Introduction

In our previous discussion, we introduced the concepts of span and linear dependence. In a way a set of vectors $S = \{v_1, v_2, \ldots, v_k\}$ span a vector space V if there are enough of the right vectors in S, while they are linearly independent if there are no redundancies. We now combine the two concepts and shall give a way to measure the size of a vector space.

Objectives
At the end of this lecture you should be able to do the following:
- study properties of elements of basis for vectors space and solve related problems,
- determine whether or not a subset of a vector space form a basis, and
- determine the number of elements in a basis for a vector space.

Pre-Test

1. Determine which of the following are bases for $\mathbb{R}^3$
   (a) $\{(1,1,1),(1,2,3),(2,-1,1)\}$
   (b) $\{(1,1,2),(1,2,5),(5,3,4)\}$.
2. Extend the linearly independent set
   $$\{(1,-1,1,-1),(1,1,-1,1)\}$$
   to a basis for $\mathbb{R}^4$.
3. Prove that $\{(3-i,2+2i,4),\ (2,2+4i,3),\ (1-i,-2i,-1)\}$ is a basis of the $\mathbb{C}$-vector space $\mathbb{C}^3$. Express each of $(1,0,0),(0,1,0),(0,0,1)$ as a linear combination of these basis vectors.

4. Let V be a vector space and let X, Y be subspaces of V. Suppose that $\dim V = 10$, $\dim X = 8$ and $\dim Y = 9$. What are the possible values of $\dim(X \cap Y)$?

5. Find a basis for the subspace of $\mathbb{R}^4$ spanned by the vectors $(1, 2, -1, 0)$, $(4, 8, -4, -3)$, $(0, 1, 3, 4)$, $(2, 5, 1, 4)$.

6. Let $W \subset \mathbb{R}^4$ be the space of solutions of the system of linear equations $AX = 0$, where $A = \begin{bmatrix} 2 & 1 & 2 & 3 \\ 1 & 1 & 3 & 0 \end{bmatrix}$. Find a basis for W.

7. How many elements are there in the vector space $\mathbb{F}_p^n$?

Definition (Basis)

Let V be a vector space and $S = \{v_1, v_2, \ldots, v_k\}$ be a subset of V. Then S is a basis for V if the following two statements are true.

1. S spans V.

2. S is a linearly independent set of vectors in V.

Example

Let $V = \mathbb{R}^n$ and let $S = \{e_1, e_2, \ldots, e_n\}$ where e_i has ith component equal to 1 and the rest 0. For example,

$$e_2 = (0, 1, 0, 0, \ldots, 0).$$

Then S is a basis for V called the standard basis.

Example

Let $V = P_3$ and let $S = \{1, t, t^2, t^3\}$. Show that S is a basis for V.

Solution

We must show both linear independence and span.

Linear Independence:

Let

$$c_1(1) + c_2(t) + c_3(t^2) + c_4(t^3) = 0$$

Then since a polynomial is zero if and only if its coefficients are all zero we have

$$c_1 = c_2 = c_3 = c_4 = 0$$

Hence S is a linearly independent set of vectors in V.

Span:

A general vector in P_3 is given by

$$a + bt + ct^2 + dt^3$$

We need to find constants c_1, c_2, c_3, c_4 such that

$$c_1(1) + c_2(t) + c_3(t^2) + c_4(t^3) = a + bt + ct^2 + dt^3$$

We just let $c_1 = a$, $c_2 = b$, $c_3 = c$, $c_4 = d$.
Hence S spans V.
We can conclude that S is a basis for V.
In general the basis $\{1, t, t^2, \ldots, t^n\}$ is called the standard basis for P_n.

Example

Show that $S = \{v_1, v_2, v_3, v_4\}$ where

$$v_1 = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix}$$

is a basis for $V = M_{2\times 2}$.

Solution

We need to show that S spans V and is linearly independent.

Linear Independence

We write

$$c_1 \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} + c_2 \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} + c_3 \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} + c_4 \begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix} = 0$$

$$\begin{pmatrix} c_1 & 2c_1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} c_2 & 0 \\ 2c_2 & 0 \end{pmatrix} + \begin{pmatrix} c_3 & 0 \\ 0 & 2c_3 \end{pmatrix} + \begin{pmatrix} 2c_4 & c_4 \\ 0 & 0 \end{pmatrix} = 0$$

this gives the four equations.

$$\begin{array}{ccccccccc} c_1 & + & c_2 & + & c_3 & + & 2c_4 & = & 0 \\ 2c_1 & & & & & + & c_4 & = & 0 \\ & & 2c_2 & & & & & = & 0 \\ & & & & 2c_3 & & & = & 0 \end{array}$$

Which has the corresponding homogeneous matrix equation

$$Ac = 0$$

with

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

We have

$$\det A = -12.$$

Since the determinant is nonzero, we can conclude that only the trivial solution exists.
That is,

$$c_1 = c_2 = c_3 = c_4 = 0.$$

Span

We set

$$c_1 \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} + c_2 \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} + c_3 \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} + c_4 \begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$\begin{pmatrix} c_1 & 2c_1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} c_2 & 0 \\ 2c_2 & 0 \end{pmatrix} + \begin{pmatrix} c_3 & 0 \\ 0 & 2c_3 \end{pmatrix} + \begin{pmatrix} 2c_4 & c_4 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

which gives the equations

$$\begin{array}{ccccccccc} c_1 & + & c_2 & + & c_3 & + & 2c_4 & = & x_1 \\ 2c_1 & & & & & + & c_4 & = & x_2 \\ & & 2c_2 & & & & & = & x_3 \\ & & & & 2c_3 & & & = & x_4 \end{array}$$

The corresponding matrix equation is

$$Ac = x$$

Since A is nonsingular, this has a unique solution, namely

$$c = A^{-1}x.$$

Hence S spans V.
We conclude that S is a basis for V.
If S spans V then we know that every vector in V can be written as a linear combination of vectors in S. If S is a basis, even more is true.

Theorem(Unique Linear Combination of Vectors)
Let $S = \{v_1, v_2, \ldots, v_k\}$ be a basis for V. Then every vector in V can be written uniquely as a linear combination of vectors in S.

Remark:
What is new here is the word uniquely.

Proof
Suppose that

$$v = a_1v_1 + \cdots + a_nv_n = b_1v_1 + \cdots + b_nv_n$$

then

$$(a_1 - b_1)v_1 + \cdots + (a_n - b_n)v_n = 0.$$

Since $S$is a basis for V, it is linearly independent and the above equation implies that all the coefficients are zero. That is

$$a_1 - b_1 = \cdots = a_n - b_n = 0.$$

We can conclude that

$$a_1 = b_1, \ldots, a_n = b_n$$

Since linear independence is all about having no redundant vectors, it should be no surprise that if $S = \{v_1, v_2, \ldots, v_k\}$ spans V, then if S is not linearly

independent then we can remove the redundant vectors until we arrive at a basis. That is, if S is not linearly independent, then one of the vectors is a linear combination of the rest. Without loss of generality, we can assume that this is the vector v_k.
We have that

$$v_k = a_1v_1 + \cdots + a_{k-1}v_{k-1}$$

If v is any vector in S we can write

$$\begin{aligned} v &= c_1v_1 + \cdots + c_kv_k \\ &= c_1v_1 + \cdots + c_{k-1}v_{k-1} + c_k(a_1v_1 + \cdots + a_{k-1}v_{k-1}) \end{aligned}$$

which is a linear combination of the smaller set.

$$S' = \{v_1, v_2, \ldots, v_{k-1}\}.$$

If S' is not a basis, as above we can get rid of another vector. We can continue this process until the vectors are finally linear independent. We have proved the following theorem.

Theorem(Subset that is a Basis for a Vector Space)
Let S span a vector space V, then there is a subset of S that is a basis for V.

Dimension

We have seen that any vector space that contains at least two vectors contains infinitely many. It is uninteresting to ask how many vectors there are in a vector space. However, there is still a way to measure the size of a vector space. For example, $\mathbb{R}^3$ should be larger than $\mathbb{R}^2$. We call this size the dimension of the vector space and define it as the number of vectors that are needed to form a basis. To show that the dimension is well defined, we need the following theorem.

Theorem
If $S = \{v_1, v_2, \ldots, v_n\}$ is a basis for a vector space V and $T = \{w_1, w_2, \ldots, w_k\}$ is a linearly independent set of vectors in V, then $k \leq n$.

Remark

If S and T are both bases for Vthen $k = n$. This says that every basis has the same number of vectors. Hence the dimension is well defined. The dimension of a vector space V is the number of vectors in a basis. If there is no finite basis we call V an infinite dimensional vector space. Otherwise, we call V a finite dimensional vector space.

Proof

If $k > n$, then we consider the set

$$R_1 = \{w_1, v_1, v_2, \ldots, v_n\}$$

Since S spans V, w_1 can be written as a linear combination of the v_i's.

$$w_1 = c_1 v_1 + \cdots + c_n v_n$$

Since T is linearly independent, w_1 is nonzero and at least one of the coefficients c_i is nonzero. Without loss of generality assume it is c_1. We can solve for v_1 and write v_1 as a linear combination of $w_1, v_2, \ldots, v_n$. Hence

$$T_1 = \{w_1, v_2, \ldots, v_n\}$$

is a basis for V. Now let

$$R_2 = \{w_1, w_2, v_2, \ldots, v_n\}.$$

Similarly, w_2 can be written as a linear combination of the rest and one of the coefficients is nonzero. Note that since w_1 and w_2 are linearly independent, at least one of the v_i coefficients must be nonzero. We can assume that this nonzero coefficient is v_2 and as before we see that

$$T_2 = \{w_1, w_2, v_3, \ldots, v_n\}$$

is a basis for V. Continuing this process we see that

$$T_n = \{w_1, w_2, \ldots, w_n\}$$

is a basis for V. But then T_n spans V and hence w_{n+1} is a linear combination of vectors in T_n. This is a contradiction since the w's are linearly independent.

Hence $n \geq k$. $\square$

Example

Since

$$E = \{e_1, e_2, \ldots, e_n\}$$

is a basis for $\mathbb{R}^n$ then $\dim \mathbb{R}^n = n$.

Example

$$\dim P_n = n + 1$$

since

$$E = \{1, t, t^2, \ldots, t^n\}$$

is a basis for P_n.

Example

$$\dim M_{m \times n} = mn$$

We will leave it to you to find a basis containing mn vectors.
If we have a set of linearly independent vectors

$$S = \{v_1, v_2, \ldots, v_k\}$$

with $k < n$, then S is not a basis. From the definition of basis, S does not span V, hence there is a v_{k+1} such that v_{k+1} is not in the span of S. Let

$$S_1 = \{v_1, v_2, \ldots, v_k, v_{k+1}\}$$

S_1 is linearly independent. We can continue this until we get a basis. Let

$$S = \{v_1, v_2, \ldots, v_k\}$$

be a linearly independent set of vectors in a vector space V. Then there are vectors $v_{k+1}, \ldots, v_n$ such that

$$\{v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_n\}$$

is a basis for V.

We finish this discussion with some very good news. We have seen that to find out if a set is a basis for a vector space, we need to check for both linearly independence and span. We know that if there are not the right number of vectors in a set, then the set cannot form a basis. If the number is the right number we have the following theorem.

Theorem(Equivalence of Basis, Linearly Independence and Spans for a Vector Space)

Let V be an n dimensional vector space and let S be a set with n vectors. Then the following are equivalent.

1. S is a basis for V.
2. S is linearly independent
3. S spans V.

Summary
We considered basis as a linearly independent spanning set. We then gave some examples, results and their respective proofs. Finally, we introduced the concept called dimension of a vector space which is defined as the number of elements in a basis.

Post-Test

See Pre-Test at the beginning of the lecture.

Supplementary Reading

1. M. Artin, Algebra: Vectors Spaces (Chapter 3), Prentice-Hall, Upper Saddle River, New Jersey 07458, 1991.
2. S.A. Ilori and O. Akinyele, Elementary Abstract and Linear Algebra, Ibadan University Press, 1986. pp. 329-371.
3. E.H. Connell, Elements of Abstract and Linear Algebra. (http://www/math.miami.edu/∼ec/book/)

4. J. Hefferon, Linear Algebra (http://joshua,smcvt.edu/linalg.html/) excellent textbook with complete solutions manual.

LECTURE 9

Coordinates and Matrices under Change of Basis

Introduction

We shall investigate how to use the coordinates from one basis into coordinates from another basis. We shall then consider definition with examples of transition matrix.

Objectives
At the end of this lecture you should be able to do the following
- solve problems on coordinates from one basis into another basis and
- define transition matrix and solve related problems.

Pre-Test

1. Let $S = \{t^2, t^2 + 2t, t^2 + 3\}$ and $T = \{2t - 1, 5t - 3, t^2\}$ be subsets of P_2
   (a) Prove that S is a basis for P_2.
   (b) Find the transition matrix $P_{S \leftarrow T}$
2. Find the representation of the polynomial
$$\begin{aligned} P(x) &= x^3 - 4x^2 + 2 \text{ in the basis } \{u_1, u_2, u_3, u_4\} \\ &= \{x^3 - 1, -x^3 + x + 1, x^3 - 2x^2, 2x^3 + x - 2\}. \end{aligned}$$
3. What is change of the basis matrix from the basis $\{v_j\} = \{2, x+1, x^2 - 2x, x^3 - x - 1\}$ to $\{w_j\} = \{1, x, x^2, x^3\}$ in P_3?
   Use it to rewrite the polynomial
$$P(x) = 3.2 + 2(x + 1) - (x^2 - 2x) + 5(x^3 - x - 1)$$

(written in the $\{v\}$ basis) into the$\{w\}$ basis.

4. What is change of basis matrix from the basis$\{w_j\} = \{1, x, x^2, x^3\}$ to the basis $\{v_j\} = \{2, x+1, x^2-2x, x^3-x-1\}$?
   Use this matrix to rewrite the polynomial

$$q(x) = 2 - 3x + 4x^2 + x^3$$

   (written in the $\{w\}$ basis) into the$\{v\}$ basis.

5. Let

$$\{v_1, v_2, v_3\} = \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} \right\}$$

   and

$$\{w_1, w_2, w_3\} = \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \right\}$$

   be two bases in $\mathbb{R}^n$. Consider the vector

$$u = 2v_1 + v_2 - 3v_3$$

   written in the $\{v\}$ basis. Find the coordinates of this vector in the $\{w\}$ basis.

Coordinates

Consider the vector $v = (2, 5, 3)$ in $\mathbb{R}^3$.
In writing these coordinates we mean

$$v = 2e_1 + 5e_2 + 3e_3$$

where

$$e_1 = (1, 0, 0),\ e_2 = (0, 1, 0),\ e_3 = (0, 0, 1)$$

are the standard basis vectors. Sometimes we are interested in finding the coordinates with respect to another basis.

Definition

Let $S = \{v_1, v_2, \ldots, v_n\}$ be a basis for a vector space V and let v be a vector in V and let

$$v = c_1 v_1 + \cdots + c_n v_n$$

Then the coordinates of v with respect to the basis S is given by

$$[v]_S = (c_1, \ldots, c_n)$$

Example

Consider the basis $S = \{(1,2),(4,7)\}$ of $\mathbb{R}^2$ and let $v = (5,8)$ presented in the standard basis. Find the coordinates of v in the basis S, that is find $[v]_S$.

Solution

We set

$$(5,8) = c_1(1,2) + c_2(4,7)$$

or

$$\begin{aligned} c_1 + 4c_2 &= 5 \\ 2c_1 + 7c_2 &= 8 \end{aligned}$$

We get the matrix equation

$$\begin{pmatrix} 1 & 4 \\ 2 & 7 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 8 \end{pmatrix}$$

Notice that the matrix is just the matrix whose columns are the basis vectors of S. The solution to this is

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 2 & 7 \end{pmatrix}^{-1} \begin{pmatrix} 5 \\ 8 \end{pmatrix} = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$$

or

$$c_1 = 3, \quad c_2 = 2.$$

What we have seen here generalizes

Theorem: Let $S = \{v_1, v_2, \ldots, v_n\}$ be a basis for a vector space V and let v be a vector in V.

Then $[v]_S = A^{-1}v$ where A is the matrix whose column vectors are$\{v_1, v_2, \ldots, v_n\}$.

The proof involves going over the previous example and generalizing. It is also interesting to run this process in reverse. $S = \{v_1, v_2, \ldots, v_n\}$ be a basis for V and let $[v]_S$ be given.
We ask how to present v in the standard basis. This follows form the theorem. Since

$$[v]_S = A^{-1}v$$

we have $v = A[v]_S$.

Example

Let $S = \{(1,3,4),(2,-1,1),(1,0,2)\}$ be a basis for $\mathbb{R}^3$ and let $[v]_S = (2,3,-1)$. Find the coordinates with respect to the standard basis.

Solution

We just find

$$\begin{pmatrix} 1 & 2 & 1 \\ 3 & -1 & 0 \\ 4 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \begin{pmatrix} 7 \\ 3 \\ 9 \end{pmatrix}$$

so that

$$v = (7,3,9).$$

Example

Let $S = \{(2,3),(1,4)\}$ and$T = \{(0,2),(-1,5)\}$ be two bases for$\mathbb{R}^2$, and let $[v]_S = (-2,6)$. Find $[v]_T$.

Solution

We can first find v in the standard basis. We have $A_S[v]_S$ where A_S is the matrix whose columns are the vectors in S. Now convert to the T basis

$$[v]_T = (A_T)^{-1}v = (A_T)^{-1}A_S[v]_S$$

or

$$[v]_T = \begin{pmatrix} 0 & -1 \\ 2 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} -2 \\ 6 \end{pmatrix} = \begin{pmatrix} 14 \\ -2 \end{pmatrix}$$

Example

Consider the vector $v = 2 + 3t - t^2$ and let $S = \{t, t-1, t^2-1\}$. Find $[v]_S$.

Solution

This problem looks a lot different from the previous ones, but looks can be deceiving. We notice that

$$\begin{aligned} t &= 0(1) + 1(t) + 0(t^2) \\ t-1 &= -1(1) + 1(t) + 0(t^2) \\ t^2-1 &= -1(1) + 0(t) + 1(t^2) \end{aligned}$$

We write that

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

and use

$$[v]_S = \begin{pmatrix} 0 & -1 & -1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \\ -1 \end{pmatrix}$$

Transition Matrices

We have seen how to use the coordinates from one basis S into coordinates from another basis T. We have

$$[v]_T = (A_T)^{-1} A_S [v]_S$$

The matrix given by

$$P_{T \leftarrow S} = (A_T)^{-1} A_S$$

is called the transition matrix from the S basis to the T basis. Note that the transition matrix from the T basis to the S basis is given by

$$P_{S \leftarrow T} = (A_S)^{-1} A_T = P_{T \leftarrow S}^{-1}$$

Example

Find the transition matrix $P_{S \leftarrow T}$ for the bases of $M^{2\times 2}$ given by

$$S = \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \right\}$$

$$T = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\}$$

Then use this matrix to find $[v]_S$ if

$$[v]_T = (1, 3, -2, 4)$$

Solution

First we denote the standard basis by

$$E = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\}$$

The A_S is just the matrix of column vectors where each column is read as you would read the matrices S. That is,

$$A_S = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

and similarly we have

$$A_T = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

The transition matrix is

$$P_{S \leftarrow T} = A_S^{-1} A_T = \begin{pmatrix} 1 & 2 & 1 & -1 \\ 0 & -1 & 0 & 1 \\ -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Now to find the coordinate in the S basis given T basis coordinates

$$[v]_T = (1, 3, -2, 4)$$

we just multiply

$$[v]_S = P_{S\leftarrow T}[v]_T = \begin{pmatrix} 1 & 2 & 1 & -1 \\ 0 & -1 & 0 & 1 \\ -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ -2 \\ 4 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 1 \\ 0 \\ 4 \end{pmatrix}$$

Remark

The transition matrix will always be nonsingular because of the nonsingular equivalence and that S and T are linearly independent.

Summary

Coordinates are defined from one basis into another basis together with their usage. We also investigate the usage of transition matrices

Post-Test

See Pre-Test at the beginning of the lecture.

Supplementary Reading

1. M. Artin, Algebra: Algebra: Vector Spaces. (Chapter 3). Prentice-Hall, Upper Saddle River, New Jersey 07458, 1991, pp 94-99.

2. D.C. Lay, Linear Algebra and Its Applications (3rd ed.), Addision Wesley, ISBN 978-0321287137 (August 22, 2005).

3. J. Hefferon, Linear Algebra (http://joshua,smcvt.edu/linalg.html/) excellent textbook with complete solutions manual.

LECTURE 10

Linear Transformations and Their Representation by Matrices

Introduction

We shall consider linear transformation and its associated properties. We shall then discuss its matrix representation.

Objectives
At the end of this lecture you should be able to do the following:
• know properties associated with linear transformation and solve related problems and
• determine, relative to ordered bases, the matrix representation of any given linear transformation

Pre-Test

1. Determine which of the following mappings $f : \mathbb{R}^3 \to \mathbb{R}^3$ are linear:
   (a) $f(x, y, z) = (z, -y, x)$;
   (b) $f(x, y, z) = (|x|, 0, -y)$;
   (c) $f(x, y, z) = (y, z, 0)$;
   (d) $f(x, y, z) = (x - 1, x, y)$;

2(a) If $f : \mathbb{R}^3 \to \mathbb{R}^3$ is linear and such that $f(1, 1, 1) = (1, 1, 1), f(1, 2, 3) = (-1, -2, -3), f(1, 1, 2) = (2, 2, 4)$ it it possible to find $f(a, b, c)$ for all $(a, b, c) \in \mathbb{R}^3$?

(b) If $f : \mathbb{R}^3 \to \mathbb{R}^3$ is linear and such that $f(1, 1, 1) = (1, 1, 1), f(2, 2, 3) = (3, 3, 5), f(1, 1, 2) = (2, 2, 4)$. Is it possible to find $f(a, b, c)$ for all $(a, b, c) \in \mathbb{R}^3$?

(c) Does there exist a linear mapping $f : \mathbb{R}^3 \to \mathbb{R}^3$ with the property that $f(0,1,1) = (3,1,-2)$, $f(1,0,1) = (4,-1,1)$, $f(1,1,0) = (-3,2,1)$, $f(1,1,1) = (3,4,2)$?

3. Find the matrices A and B associated with the linear mappings. $f : \mathbb{R}^2 \to \mathbb{R}^3$ and $g : \mathbb{R}^3 \to \mathbb{R}^2$ with respect to the standard bases when f and g are given by $f(x,y) = (x+2y, 2x-y, -x)$, $g(x,y,z) = (2x-y, 2y-z)$. Find the matrices C and D associated with these mappings f and g with respect to the bases $\{(0,1),(1,1)\}$ of $\mathbb{R}^2$ and $\{(0,0,1),(0,1,1),(1,1,1)\}$ of $\mathbb{R}^3$.

4. Let L be the linear transformation from $\mathbb{R}^2$ toP_2 defined by

$$L(x,y) = x + yt + (x+y)t^2$$

Find the matrix representing L with respect to the standard bases.

5. Let L be the linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ such that

$$L(x,y) = (x-2y, y-2x)$$

and let

$$S = \{(2,3),(1,2)\}$$

be a basis for$\mathbb{R}^2$. Find the matrix for L that sends a vector from the S basis to the standard basis.

6. Let L be the linear transformation from P_2 to P_2 with such that

$$L(a+bt+ct^2) = (a+c) + (a+2b)t + (b+3c)t^2$$

and let

$$S = (1-t, 1-t^2, t-t^2) \text{ and } T = (2+t+t^2, 1+t, 1+t+t^2).$$

Find the matrix of L with respect to the bases S and T.

7. Let L be the linear transformation from $M_{2\times 2}$ to $M_{2\times 2}$ and let

$$L\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & a \\ b & c \end{pmatrix}$$

and

$$S = \left\{ \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \right\}$$

Find the matrix for L from S to S.

Linear Transformations

A linear transformation is a function from one vector space to another that preserves the vector space operations.
Let us be more precise.

Definition

Let V and W be vector spaces over a field F. A function $\tau : V \to W$ is a linear transformation if

$$\tau(ru + sv) = r\tau(u) + s\tau(v)$$

for all scalars $r, s \in F$ and vectors $u, v \in V$.
A linear transformation $\tau : V \to V$ is called a linear operator on V.
The set of all linear transformations from V to W is denoted by $\mathcal{L}(V, W)$ and the set of all linear operators on V is denoted by $\mathcal{L}(V)$.

Definition

The following terms are also employed

(1) homomorphism for linear transformation

(2) endomorphism for linear operator

(3) monomorphism(or embedding) for injective linear transformation

(4) epimorphism for surjective linear transformation

(5) isomorphism for bijective linear transformation

(6) automorphism for bijective linear operator.

Example

1. The derivative $D : V \to V$ is a linear operator on the vector space V of all infinitely differentiable functions on $\mathbb{R}$.
2. The integral operator $\tau : F[x] \to F[x]$ defined by
$$\tau(f) = \int_0^x f(t)dt$$
is a linear operator on $F[x]$.
3. Let A be an $m \times n$ matrix over F. The function $\tau_A : F^n \to F^m$ defined by $\tau_A(v) = Av$, where all vectors a written as column vectors, is a linear transformation from F^n to F^m. This function is just multiplication by A.

Theorem

1. The set $\mathcal{L}(V, W)$ is a vector space under ordinary addition of functions and scalars multiplication of functions by elements of F.
2. If $\sigma \in \mathcal{L}(U, V)$ and $\tau \in \mathcal{L}(V, W)$ then the composition $\tau\sigma$ is in $\mathcal{L}(U, W)$.
3. If $\tau \in \mathcal{L}(V, W)$ is bijective then $\tau^{-1} \in \mathcal{L}(W, V)$.
4. The vector space $\mathcal{L}(V)$ is an algebra, where multiplication is composition of functions. The identity map $i \in \mathcal{L}(V)$ is the multiplication identity and the zero map $o \in \mathcal{L}(V)$ is the additive identity.

Proof

We prove only Part (3).

Let $\tau : V \to W$ be a bijective linear transformation. Then $\tau^{-1} : W \to V$ is a well-defined function and since any two vectors w_1 and w_2 in W have the form $w_1 = \tau(v_1)$ and $w_2 = \tau(v_2)$, we have

$$\begin{aligned} \tau^{-1}(aw_1 + bw_2) &= \tau^{-1}(a\tau(v_1) + b\tau(v_2)) \\ &= \tau^{-1}(\tau(av_1 + bv_2)) \\ &= av_1 + bv_2 \\ &= a\tau^{-1}(w_1) + b\tau^{-1}(w_2) \end{aligned}$$

which shows that τ^{-1} is linear.

One of the easiest ways to define a linear transformation is to give its values on a basis.

Theorem

Let V nd W be vector spaces and let $\mathcal{B} = \{v_i : i \in I\}$ be a basis for V. Then we can define a linear transformation $\tau \in \mathcal{L}(V, W)$ by specifying the values of $\tau(v_i) \in W$ arbitrarily for all $v_i \in \mathcal{B}$ and extending the domain of τ to V using linearity, that is,

$$\tau(a_1v_1 + \cdots + a_nv_n) = a_1\tau(v_1) + \cdots + a_n\tau(v_n)$$

This process uniquely defines a linear transformation, that is, if $\tau, \sigma \in \mathcal{L}(V, W)$ satisfy $\tau(v_i) = \sigma(v_i)$ for all $v_i \in \mathcal{B}$ then $\tau = \sigma$.
Note that if $\tau \in \mathcal{L}(V, W)$ and if S is a subspace if V, then the restriction $\tau|_S$ of τ to S is a linear transformation from S to W.

Practice Exercise

Let $L : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation such that

$$L(1, 4) = (1, -1, 3) \text{ and } L(0, 2) = (2, 1, 4)$$

Find $L(1, 0)$.

The Kernel and Image of a Linear Transformation
Definition

Let $\tau \in \mathcal{L}(V, W)$. The subspace

$$ker(\tau) = \{v \in V : \tau(v) = 0\}$$

is called the kernel of τ and the subspace

$$im(\tau) = \{\tau(v) : v \in V\}$$

is called the image of τ.
The dimension of $ker(\tau)$ is called the nullity if τ and is denoted by null(τ).
The dimension of $im(\tau)$ is called the rank of τ and is denoted by $rk(\tau)$.

Practice Exercise

Show that $ker(\tau)$ is a subspace of V and $im(\tau)$ is a subspace of W.

Theorem

Let $\tau \in \mathcal{L}(V, W)$. Then

(1) τ is surjective if and only if $im(\tau) = W$

(2) τ is injective if and only if $ker(\tau) = \{0\}$

Proof

The first statement is merely a restatement of the definition of surjectivity. To see the validity of the second statement, observe that

$$\tau(u) = \tau(v) \Leftrightarrow \tau(u - v) = 0 \Leftrightarrow u - v \in ker(\tau)$$

Hence, if $ker(\tau) = \{0\}$, then $\tau(u) = \tau(v) \Leftrightarrow u = v$, which shows that τ is injective.

Conversely,m if τ is injective and $u \in ker(\tau)$ then $\tau(u) = \tau(0)$ and so $u = 0$.
This shows that $ker(\tau) = \{0\}$.

Matrix of a Linear Transformation
Theorem

(1) If A is an $m \times n$ matrix over F then $\tau_A \in \mathcal{L}(F^n, F^m)$.

(2) If $\tau \in \mathcal{L}(F^n, F^m)$ then $\tau = \tau_A$ where

$$A = (\tau(e_1)| \cdots |\tau(e_n))$$

The matrix A is called the matrix of τ.

Example

Consider the linear transformation $\tau : F^3 \to F^3$ defined by

$$\tau(x, y, z) = (x - 2y, z, x + y + z)$$

Then we have, in column form

$$\tau \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-2y \\ z \\ x+y+z \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

and so the standard matrix of τ is

$$A = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Let $\tau : V \to W$ be a linear transformation, where $\dim(V) = n$ and $\dim(W) = m$ and let $\mathcal{B} = (b_1, \ldots, b_n)$ be an ordered basis for V and $\mathcal{L}$ an ordered basis for W. Then the map

$$\theta : [v]_{\mathcal{B}} \to [\tau(v)]_C$$

is a representation of τ as a linear transformation from F^n to F^m, in the sense that knowing θ (along with $\mathcal{B}$ and $\mathcal{C}$ of course) is equivalent to knowing τ. Of course, this representation depends on the choice of ordered bases $\mathcal{B}$ and $\mathcal{C}$.

Since θ is a linear transformation from F^n to F^m, it is just multiplication by an$m \times n$ matrix A, that is

$$[\tau(v)]_{\mathcal{C}} = A[v]_{\mathcal{B}}$$

Theorem

Let $\tau \in \mathcal{L}(V, W)$ and let $\mathcal{B} = (b_1, \ldots, b_n)$ and $\mathcal{C}$ be ordered bases for V and W, respectively. Then τ can be represented with respect to $\mathcal{B}$ and $\mathcal{C}$ as matrix multiplication, that is

$$[\tau(v)]_{\mathcal{C}} = [\tau]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}}$$

where

$$[\tau]_{\mathcal{B},\mathcal{C}} = ([\tau(b_1)]_{\mathcal{C}}| \cdots |[\tau(b_n)]_{\mathcal{C}})$$

is called the matrix of τ with respect to the bases $\mathcal{B}$ and $\mathcal{C}$. When $V = W$ and $\mathcal{B} = \mathcal{C}$, we denote $[\tau]_{\mathcal{B},\mathcal{B}}$ by $[\tau]_{\mathcal{B}}$ and so

$$[\tau(v)]_{\mathcal{B}} = [\tau]_{\mathcal{B}}[v]_{\mathcal{B}}$$

Example

Let $D : \mathcal{P}_2 \to \mathcal{P}_2$ be the derivative operator, defined on the vector space of all polynomials of degree at most 2. Let $\mathcal{B} = \mathcal{C} = (1, x, x^2)$. Then

$$[D(1)]_{\mathcal{C}} = [0]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

$$[D(x)]_{\mathcal{C}} = [1]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

$$[D(x^2)]_{\mathcal{C}} = [2x]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

and so

$$[D]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence, for example, if $p(x) = 5 + x + 2x^2$ then

$$\begin{aligned} [D(p(x))]_{\mathcal{C}} &= [D]_{\mathcal{B}}[p(x)]_{\mathcal{B}} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix} \end{aligned}$$

and so $Dp(x) = 1 + 4x$.

Summary

We introduce linear transformations and their properties , proofs and examples We considered how to find the matrix for a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$. Finally, we showed how to find the matrix of a general linear transformation when the bases are given

Post-Test

See Pre-Test at the beginning of the lecture.

Supplementary Reading

1. M. Artin, Algebra: Matrix Operations (Chapter 1). Prentice-Hall, Upper Saddle River, New Jersey 07458, 1991, pp 111-115.

2. D.C. Lay, Linear Algebra and Its Applications (3rd ed.), Addision Wesley, ISBN 978-0321287137 (August 22, 2005).

3. J. Hefferon, Linear Algebra (http://joshua,smcvt.edu/linalg.html/) excellent textbook with complete solutions manual.

LECTURE 11

Eigenvalues and Eigenvectors

Introduction

If L is a linear transformation from a vector space to itself, then of interest is whether there are any vectors v having the property that $L(v)$ is a multiple of v. If this is the case then repeatedly applying L will result in a vector always parallel to v. This idea is of fundamental importance for applications in physics, mechanics, economics, biology, and just about every other scientific field.

Objectives
At the end of this lecture, you should be able to do the following:
• use characteristic equations and polynomials to determine eigenvalues and eigenvectors of a given square matrix,
• compute a huge power of a square matrix,
• use principle of diagonalization eigenvalues and eigenvectors properties to solve a given system of linear differential equations, and
• extend the above ideas to $n \times n$ matrices.

Pre-Test

1. Let

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}$$

(a) Verify that $\det(\lambda I_3 - A)$, the characteristic polynomial of A, is given by

$$(\lambda - 1)\lambda(\lambda - \frac{1}{4})$$

(b) Find a non-singular matrix P such that

$$P^{-1}AP = diag(1, 0, \frac{1}{4}).$$

(c) Prove that

$$A^n = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} + \frac{1}{3 \cdot 4^n}\begin{bmatrix} 2 & 2 & -4 \\ -1 & -1 & 2 \\ -1 & -1 & 2 \end{bmatrix}$$

if $n \geq 1$.

2. Let

$$A = \begin{bmatrix} 5 & 2 & -2 \\ 2 & 5 & -2 \\ -2 & -2 & 5 \end{bmatrix}$$

(a) Verify that $\det(\lambda I_3 - A)$, the characteristic polynomial of A, is given by

$$(\lambda - 3)^2(\lambda - 9)$$

(b) Find a non-singular matrix P such that

$$P^{-1}AP = diag(3, 3, 9).$$

3. Solve the system of differential equations

$$\frac{dx}{dt} = 3x - 2y$$

$$\frac{dy}{dt} = 5x - 4y$$

given $x = 13$ and $y = 22$ when $t = 0$.

4. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a real or complex matrix with distinct eigenvalues λ_1, λ_2 and corresponding eigenvectors X_1, X_2.

Prove that the system of differential equations

$$\frac{dx}{dt} = ax + by$$

$$\frac{dy}{dt} = cx + dy$$

has the solution

$$\begin{bmatrix} x \\ y \end{bmatrix} = \alpha e^{\lambda_1 t} X_1 + \beta e^{\lambda_2 t} X_2.$$

where α and β are determined by the equation

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = P^{-1} \begin{bmatrix} x(0) \\ y(0) \end{bmatrix}$$
$$P = [X_1 | X_2].$$

5. For each of the matrices

$$\begin{bmatrix} -3 & -7 & 19 \\ -2 & -1 & 8 \\ -2 & -3 & 10 \end{bmatrix}, \quad \begin{bmatrix} -4 & 0 & -3 \\ 1 & 3 & 1 \\ 4 & -2 & 2 \end{bmatrix}$$

Find a matrix T such that $T^{-1}AT$ is diagonal.

6. For every positive integer n, determine the nth power of he matrix.

$$A = \begin{bmatrix} 2 & 2 & 0 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix}$$

Definition (Eigenvector and Eigenvalue)

Let A be a complex square matrix. Then if λ is a complex number and v a non-zero complex column vector satisfying $Av = \lambda v$, we call v an eigenvector of A, while λ is called an eigenvalue of A. We also say that v is an eigenvector corresponding to the eigenvalue λ.

Finding Eigenvalues and Eigenvectors

We will now determine how to find the eigenvalues and eigenvectors for a matrix. Let A be a matrix with eigenvalue λ and eigenvector v. Then

$$Av = \lambda v$$

implies that

$$Av - \lambda v = 0$$

We would like to factor out a v from the above (right) equation. However, we must be careful because the term $A - \lambda$ does not make sense. Instead we have

$$(A - \lambda I)v = 0$$

If there is a nontrivial solution, then

$$\det(A - \lambda I) = 0$$

and the solution is in the null space of

$$A = \lambda I.$$

Example

Find the eigenvalues and eigenvectors of

$$A = \begin{pmatrix} 3 & 2 \\ 3 & 4 \end{pmatrix}$$

we have

$$A - \lambda I = \begin{pmatrix} 3-\lambda & 2 \\ 3 & 4-\lambda \end{pmatrix}$$

which has determinant

$$\begin{aligned} (3-\lambda)(4-\lambda) - 6 &= \lambda^2 - 7\lambda + 12 - 6 \\ &= \lambda^2 - 7\lambda + 6 \\ &= (\lambda - 1)(\lambda - 6) \end{aligned}$$

So the roots are

$$\lambda = 1 \text{ and } \lambda = 6.$$

Now, let's find the eigenvectors.
For $\lambda = 1$, we have

$$A - I = \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix}$$

which has row reduced echelon form

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

A nonzero vector in the null space is

$$V_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Now for the eigenvector corresponding to $\lambda = 6$, we have

$$A - 6I = \begin{pmatrix} -3 & 2 \\ 3 & -2 \end{pmatrix}$$

Notice that the second row is redundant. At this point, it is pretty easy to see that

$$V_6 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

Notice that there are lots of choices for these vectors (all multiples of the above vector). We made our choice in order to avoid fractions.

Definition (Characteristic Polynomial and Equation of a Matrix)

The equation $\det(A - \lambda I_n) = 0$ is called the characteristic equation of A, while the polynomial $\det(A - \lambda I_n)$ is called the characteristic polynomial of A. The characteristic polynomial of A is often denoted by $ch_A(\lambda)$.

Hence the eigenvalues of A are the roots of the characteristic polynomial of A.

For a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, it is easily verified that the characteristic polynomial is $\lambda^2 - (\text{trace } A)\lambda + \det A$, where trace $A = a + d$ is the sum of the diagonal elements of A.

Example

Find the eigenvalues of $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, and find all eigenvectors.

Solution

The characteristic equation of A is

$$\lambda^2 - 4\lambda + 3 = 0,$$

or

$$(\lambda - 1)(\lambda - 3) = 0$$

Hence $\lambda = 1$ or 3. The eigenvector equation $(A - \lambda I_n)v = 0$ reduces to

$$\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or

$$\begin{aligned} (2-\lambda)x &+ y &= 0 \\ x &+ (2-\lambda)y &= 0 \end{aligned}$$

Taking $\lambda = 1$ gives

$$\begin{aligned} x + y &= 0 \\ x + y &= 0, \end{aligned}$$

which has solution $x = y$, y arbitrary.

Consequently, the eigenvectors corresponding to $\lambda = 1$ are the vectors $\begin{bmatrix} -y \\ y \end{bmatrix}$, with $y \neq 0$.

Taking $\lambda = 3$ gives

$$\begin{aligned} -x + y &= 0, \\ x - y &= 0, \end{aligned}$$

which has solution $x = y$, y arbitrary.

Consequently, the eigenvectors corresponding to $\lambda = 3$ are the vectors $\begin{bmatrix} y \\ y \end{bmatrix}$,

with $y \neq 0$.
Our next result has wide applicability:

Theorem

Let A be a 2×2 matrix having distinct eigenvalues λ_1 and λ_2 and corresponding eigenvectors v_1 and v_2. Let P be the matrix whose columns are v_1 and v_2, respectively.
Then P is non-singular and

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Example

Let $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ be the matrix of above example. Then $v_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ are eigenvectors corresponding to eigenvalues 1 and 3, respectively.
Hence, if

$$P = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix},$$

we have

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

There are two immediate applications of above theorem. The first is to the calculation of A^n: If $P^{-1}AP = diag(\lambda_1, \lambda_2)$, then $A = P\, diag(\lambda_1, \lambda_2)P^{-1}$ and

$$\begin{aligned} A^n &= \left(P \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} P^{-1}\right)^n = P \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}^n P^{-1} \\ &= P \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} P^{-1} \end{aligned}$$

The second application is to solving a system of linear differential equations

$$\frac{dx}{dt} = ax + by$$

$$\frac{dy}{dt} = cx + dy,$$

where

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is a matrix of real or complex numbers and x and y are functions of t. The system can be written in matrix form as $\dot{v} = Av$, where

$$v = \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{and} \quad \dot{v} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} \frac{dx}{dt} \\ \\ \frac{dy}{dt} \end{bmatrix}$$

We make the substitution $v = Pw$, where $w = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$. Then, x_1 and y_1 are also functions of t and

$$\dot{v} = P\dot{w} = Av = A(Pw),$$

so

$$\dot{w} = (P^{-1}AP)w = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} w.$$

Hence $\dot{x}_1 = \lambda_1 x_1$ and $\dot{y}_1 = \lambda_2 y_1$.
These differential equations are well-known to have the solutions $x_1 = x_1(0)e^{\lambda_1 t}$ and $y_1 = y_1(0)e^{\lambda_2 t}$, where $x_1(0)$ is the value of x_1 when $t = 0$.
[If $\dfrac{dx}{dt} = kx$, where k is a constant, then

$$\begin{aligned} \frac{d}{dt}(e^{-kt}x) &= -ke^{-kt}x + e^{-kt}\frac{dx}{dt} \\ &= -ke^{-kt}x + e^{-kt}kx = 0 \end{aligned}$$

Hence $e^{-kt}x$ is constant, so $e^{-kt}x = e^{-k0}x(0) = x(0)$.
Hence $x = x(0)e^{kt}$].
However,

$$\begin{bmatrix} x_1(0) \\ y_1(0) \end{bmatrix} = P^{-1} \begin{bmatrix} x(0) \\ y(0) \end{bmatrix},$$

so this determines $x_1(0)$ and $y_1(0)$ in terms of $x(0)$ and $y(0)$.
Hence, ultimately x and y are determined as explicit functions of t, using the equation $v = Pw$.

Example

Let $A = \begin{bmatrix} 2 & -3 \\ 4 & -5 \end{bmatrix}$. Use the eigenvalue method to derive an explicit formula for A^n and also solve the system of differential equations

$$\frac{dx}{dt} = 2x - 3y$$

$$\frac{dy}{dt} = 4x - 5y,$$

given $x = 7$ and $y = 13$ when $t = 0$.

Solution

The characteristic polynomial of A is $\lambda^2 + 3\lambda + 2$ which has distinct roots $\lambda_1 = -1$ and $\lambda_2 = -2$. We find corresponding eigenvectors $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$. Hence if $P = \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix}$, we have $P^{-1}AP = diag(-1, -2)$.
Hence,

$$\begin{aligned}
A^n &= (P\ diag(-1,-2)P^{-1})^n \\
&= P\ diag((-1)^n, (-2)^n)P^{-1} \\
&= \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} (-1)^n & 0 \\ 0 & (-2)^n \end{bmatrix} \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix} \\
&= (-1)^n \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2^n \end{bmatrix} \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix} \\
&= (-1)^n \begin{bmatrix} 1 & 3 \times 2^n \\ 1 & 4 \times 2^n \end{bmatrix} \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix} \\
&= (-1)^n \begin{bmatrix} 4 - 3 \times 2^n & -3 + 3 \times 2^n \\ 4 - 4^n \times 2^n & -3 + 4 \times 2^n \end{bmatrix}.
\end{aligned}$$

To solve the differential equation system, make the substitution $v = Pw$. Then

$$x = x_1 + 3y_1, \quad y = x_1 + 4y_1.$$

The system then becomes

$$\begin{aligned} \dot{x}_1 &= -x_1 \\ \dot{y}_1 &= -2y_1, \end{aligned}$$

so, $x_1 = x_1(0)e^{-t}$, $y_1 = y_1(0)e^{-2t}$.
Now,

$$\begin{bmatrix} x_1(0) \\ y_1(0) \end{bmatrix} = P^{-1} \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 7 \\ 13 \end{bmatrix} = \begin{bmatrix} -11 \\ 6 \end{bmatrix}.$$

so,
$x_1 = -11e^{-t}$ and $y_1 = 6e^{-2t}$. Hence,

$$x = -11e^{-t} + 3(6e^{-2t}) = -11e^{-t} + 18e^{-2t},$$

$$y = -11e^{-t} + 4(6e^{-2t}) = -11e^{-t} + 24e^{-2t}.$$

Theorem(An Equivalence Result)
Let A be an $n \times n$ matrices. Then the following conditions are equivalent.

1. A is nonsingular
2. $Av = 0$ has only the trivial solution
3. A is row equivalent to I
4. $Av = b$ has a unique solution for all b.
5. $\det(A)$ is nonzero
6. A has rank n
7. A has nullity 0.
8. The rows of A are linearly independent

9. The columns of A are linearly independent

10. 0 is not an eigenvalue of A.

Summary
Eigenvalues and eigenvectors are defined.
We then applied characteristic equations and polynomials of a given square matric to solve for eigenvalues and eigenvectors. Finally we used digonalization principle to calculate huge power of a square matrix as well as solving linear system of differential equations.

Post-Test
See Pre-Test at the beginning of the lecture.

Supplementary Reading

1. M. Artin, Algebra, Linear Transformations, (Chapter 4), Prentice-Hall, Upper Saddle River, New Jersey 074548, 1991, pp. 115-138.

2. D.C. Lay, Linear Algebra and Its Applications (3rd ed.), Addision Wesley, ISBN 978-0321287137 (August 22, 2005).

3. J. Hefferon, Linear Algebra (http://joshua,smcvt.edu/linalg.html/) excellent textbook with complete solutions manual.

LECTURE 12

Diagonalisation

Introduction

We have seen that the commutation property does not hold for matrices, so that if A is an $n \times n$ matrix, then

$$P^{-1}AP$$

is not necessarily equal to A.

For different nonsingular matrices P, the above expression will represent different matrices. However, all such matrices share some important properties as we shall soon see.

Objectives
At the end of this lecture you should be able to do the following:
- define and identify classes of similar matrices,
- verify that classes of similar matrices have the same eigenvalue, and
- know when matrix is diagonalizable and (if possible) determine its similar diagonal matrix.

Pre-Test

1. In each case, if the matrix is diagonalizable, find a matrix P such that PAP^{-1} is diagonal.

(a) $\begin{bmatrix} -2 & 2 \\ -2 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$

(d) $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

2. Prove that if A, B are $n \times n$ matrices and if A is nonsingular, then AB is similar to BA.

3. Diagonalize the matrix

(a) $A = \begin{pmatrix} 2 & 5 \\ -1 & -4 \end{pmatrix}$

(b) $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

(c) $C = \begin{pmatrix} 2 & -2 & -1 \\ 0 & 1 & 0 \\ 2 & -4 & -1 \end{pmatrix}$

Definition (Similar Matrices)

Let A and B be $n \times n$ matrices, then A is similar to B if there is a nonsingular matrix P with

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

Example

Consider the matrices

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 5 \end{pmatrix}, \qquad P = \begin{pmatrix} 3 & 4 \\ 4 & 5 \end{pmatrix}$$

Then

$$\begin{aligned} B = P^{-1}AP &= \begin{pmatrix} -5 & 4 \\ 4 & -3 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} 3 & 4 \\ 4 & 5 \end{pmatrix} \\ &= \begin{pmatrix} 82 & 101 \\ -61 & -75 \end{pmatrix} \end{aligned}$$

is similar to A.
Notice the three following facts

1. A is similar to A

2. If A is similar to B then B is similar to A.

3. If A is similar to B and B is similar to C then A is similar to C.

We call a relationship with these three properties an equivalence relationship. We will prove the third property.
If A is similar to B and B is similar to C, then there are matrices P and Q with $B = P^{-1}AP$ and $C = Q^{-}BQ$. We need to find a matrix R with

$$C = R^{-1}AR.$$

We have

$$\begin{aligned} C = Q^{-1} \cdot BQ &= Q^{-1}(P^{-1}AP)Q \\ &= (Q^{-1}P^{-1})A(PQ) \\ &= (PQ)^{-1}A(PQ) \\ &= R^{-1}AR \end{aligned}$$

There is a wonderful fact that we state below.

Theorem

If A and B are similar matrices, then they have the same eigenvalues.

Proof

It is enough to show that they have the same characteristic polynomials. We have

$$\begin{aligned} \det(\lambda I - B) &= \det(\lambda I - P^{-1}AP) \\ &= \det(P^{-1}\lambda IP - P^{-1}AP) \\ &= \det(P^{-1}(\lambda I - A)P) = \det(\lambda I - A) \end{aligned}$$

Diagonalized Matrices

The easiest kind of matrices to deal with are diagonal matrices. Determinants are simple, the eigenvalues are just the diagonal entries and the eigenvectors are just the elements of the standard basis. Even the inverse is a piece of cake (if the matrix is nonsingular). Although most matrices are not diagonal, many are diagonalisable, that is they are similar to a diagonal matrix.

Definition (Diagonalisable Matrix)

A matrix A is diagonalisable if A is similar to a diagonal matrix D. i.e.,

$$D = P^{-1}AP.$$

The following theorem tells us when a matrix is diagonalisable and if it is, how to find its similar diagonal matrix D.

Theorem

Let A be an $n \times n$ matrix. Then A is diagonalisable if and only if A has n linearly independent eigenvectors. If so, then

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

If $\{v_1, \ldots, v_n\}$ are the eigenvectors of A and $\{\lambda_1, \ldots, \lambda_n\}$ are the corresponding eigenvalues, then v_j, the jth column of P and

$$[D]_{jj} = \lambda_j$$

Example

In the last discussion, we saw that the matrix

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix}$$

has -1 and 4 as eigenvalues with associated eigenvectors

$$v_{-1} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}, \qquad v_4 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Hence,

$$P = \begin{pmatrix} 3 & 1 \\ -2 & 1 \end{pmatrix}, \qquad D = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}$$

You can verify that

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

Theorem

Let A be an $n \times n$ matrix with n real and distinct eigenvalues. Then A is diagonalisable.

Note that the converse certainly does not hold. For example, the identity matrix I has 1 as all of its eigenvalues, but it is diagonalisable (it is diagonal).

Steps to Diagonalise a Matrix

1. Find the eigenvalues by finding the roots of the characteristic polynomial.
2. Find the eigenvectors by finding the nullspace of $A - \lambda_i I$.
3. If the number of linearly independent vectors is n, then let P be the matrix whose columns are eigenvectors and let D be the diagonal matrix with $[D]_{jj} = \lambda_j$.

Example

Diagonalise the matrix

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 0 & 1 & 0 \\ 2 & 1 & 0 \end{pmatrix}$$

Solution

We find the characteristic polynomial

$$\begin{aligned} \det(\lambda I - A) &= \left| \begin{pmatrix} \lambda - 3 & -1 & 1 \\ 0 & \lambda - 1 & 0 \\ -2 & -1 & \lambda \end{pmatrix} \right| \\ &= (\lambda - 3)(\lambda - 1)\lambda + 2(\lambda - 1) \\ &= (\lambda - 1)(\lambda^2 - 3\lambda + 2) \\ &= (\lambda - 1)^2(\lambda - 2) \end{aligned}$$

The roots are 1 (with multiplicity 2) and 2 (with multiplicity 1).

Now we find the eigenspaces associated with the eigenvalues. We have row reduced echelon form $((1)I - A)$

$$\begin{aligned}\text{row reduced echelon form} &= \begin{pmatrix} -2 & -1 & 1 \\ 0 & 0 & 0 \\ -2 & -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\end{aligned}$$

A basis for the null space is

$$V_1 = \left\{ \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \right\}$$

Next we find a basis for the eigenspace associated with the eigenvalue 2. We have

$$\begin{aligned} rref((2)I - A) &= rref \begin{pmatrix} -1 & -1 & 1 \\ 0 & 1 & 0 \\ -2 & -1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

A basis for this null space is

$$v_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Now put this all together to get

$$P = \begin{pmatrix} -1 & 1 & 1 \\ 2 & 0 & 0 \\ 0 & 2 & 1 \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Example

Diagonalise the matrix

$$A = \begin{pmatrix} 2 & -2 & -1 \\ 0 & 1 & 0 \\ 2 & -4 & -1 \end{pmatrix}$$

Solution

Run the usual machinery:

$$\begin{aligned} Ch_A(\lambda) &= \det \begin{pmatrix} 2-\lambda & -2 & -1 \\ 0 & 1-\lambda & 0 \\ 2 & -4 & -1-\lambda \end{pmatrix} \\ &= -\lambda(1-\lambda)^2 \end{aligned}$$

Hence $\lambda_1 = 0$, $\lambda_2 = \lambda_3 = 1$, i.e., the eigenvalue 1 has multiplicity 2.
First find the eigenvector v_1 to $\lambda_1 = 0$.
Since λ_1 is a simple root, this is easy, as before you solve

$$\begin{pmatrix} 2 & -2 & -1 \\ 0 & 1 & 0 \\ 2 & -4 & -1 \end{pmatrix} v_1 = \begin{pmatrix} 2 & -2 & -1 \\ 0 & 1 & 0 \\ 2 & -4 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and easily find (one) solution $v_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$.

Next, you try to find two linearly independent eigenvectors to $\lambda = 1$, i.e., you solve

$$\begin{pmatrix} 1 & -2 & -1 \\ 0 & 0 & 0 \\ 2 & -4 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

After Gauss,

$$\left(\begin{array}{ccc|c} 1 & -2 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & -4 & -2 & 0 \end{array}\right) \Longrightarrow \left(\begin{array}{ccc|c} 1 & -2 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

i.e. you have TWO free variables! Hence, you will have two linearly independent solutions. How to find them? Exactly as you found a basis in the

null space of a matrix: Set a free variable 1, the rest zero, and repeat for all free variables. E.g. here y, z are free, set first $y = 1$, $z = 0$, get $x = 2$, i.e. $v_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$; then set $y = 0$, $z = 1$, get $x = 1$, i.e. $v_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.
Finally, you will have to invent the matrix

$$V = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

the result is

$$V^{-1} = \begin{pmatrix} -1 & 2 & 1 \\ 0 & 1 & 0 \\ 2 & -4 & -1 \end{pmatrix}$$

and the diagonalisation of A.

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 2 & 1 \\ 0 & 1 & 0 \\ 2 & -4 & -1 \end{pmatrix}$$

Practice Exercise

Let $A = \begin{pmatrix} 2 & 5 \\ -1 & -4 \end{pmatrix}$ and let $w = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.
Compute $A^{2000}w$.

Summary
Similar Classes of Matrices are considered. We then stated steps to Diagonalise a matrix with examples to verify.

Post-Test

See Pre-Test at the beginning of the lecture.

Supplementary Reading

1. M. Artin, Algebra, Linear Transformations, (Chapter 4), Prentice-Hall, Upper Saddle River, New Jersey 074548, 1991, pp. 130-132.

2. D.C. Lay, Linear Algebra and Its Applications (3rd ed.), Addision Wesley, ISBN 978-0321287137 (August 22, 2005).

3. E.H. Connell, Elements of Abstract and Linear Algebra. (http://www/math.miami.edu/∼ec/book/)

4. J. Hefferon, Linear Algebra (http://joshua,smcvt.edu/linalg.html/) excellent textbook with complete solutions manual.

LECTURE 13

Groups

Introduction

A Group is basically a study of sets with single binary operations, respectively. We shall consider their simple properties and give examples which will include the abelian, cyclic and symmetric groups.

In studying groups, we shall also consider the two important notions: subgroups and the structure preserving mapping from one such group to another.

Objectives

At the end of this lecture you should be able to do the following:

- give different examples of binary operations,
- show whether a set together with a given binary operation forms a group,
- give different examples of a group,
- determine subgroups of a given groups, and
- give simple examples of homomorphisms of groups

Pre-Test

1. Let $P = \{p \in \mathbb{Z} : p \text{ is a prime and } p \leq 13\}$.
   Define a binary operation $*$ on P by $p * q =$ the greatest prime divisor of $p + q - 2$. Construct the Cayley table for $*$ and show that P has an identity with respect to $*$. Does every element of P have an inverse? Is $*$ associative?

2. Define a binary operation $*$ on $\mathbb{R}^* = \mathbb{R}\backslash\{0\}$ by

$$a * b = \begin{cases} ab & \text{if } a > 0; \\ \dfrac{a}{b} & \text{if } a < 0. \end{cases}$$

Is $\mathbb{R}^*$ a group under $*$?

3. The following is part of the Cayley table of a finite group. Fill in the missing entries:

	e	a	b	x	y	z
e	e	a	b	x	y	z
a	a	b	e	y		
b	b					
x	x	z				a
y	y					
z	z					

4. Show that $H = \{\bar{0}, \bar{4}, \bar{8}, \overline{12}\}$ is a subgroup of $\mathbb{Z}_{16}$ under addition.

5. If G is an abelian group and $a, b \in G$ are distinct elements of order 2, show that ab has order 2. Prove that$\{1, a, b, ab\}$ forms a subgroup of G that is not cyclic.

6(a) Write the element

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 6 & 9 & 7 & 2 & 5 & 8 & 1 & 3 \end{pmatrix}$$

of S_9 as a product of disjoint cycles

(b) Express each of the following permutations as products of disjoint cycles:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 2 & 6 & 3 & 7 & 4 & 5 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 6 & 4 & 1 & 8 & 2 & 5 & 7 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 1 & 4 & 7 & 2 & 5 & 8 & 6 \end{pmatrix}.$$

(c) Write $(143)(534)(137)$ as a product of disjoint cycles.

7. Do the elements $(12)(34)$ and $(13)(24)$ of S_4 commute? Do the elements $(12)(24)$ and $(13)(34)$ of S_4 commute?

8. Which of the following are group morphism?

   (a) $f : \mathbb{Z}_{12} \to \mathbb{Z}_{12},\ f(x \bmod 12) = (x+1) \bmod 12.$

   (b) $f : C_{12} \to C_{12},\ f(g) = g^3$

   (c) $f : \mathbb{Z} \to \mathbb{Z}_2 \times \mathbb{Z}_4,\ f(x) = (x(\text{mod}2), x(\text{mod}4))$

   (d) $f : \mathbb{Z}_8 \to \mathbb{Z}_2,\ f(x(\text{mod}8)) = x(\text{mod}2)$

   (e) $f : C_2 \times C_3 \to S_3,\ f(h^r, k^s) = (12)^2(123)^s;$

9. Show that $G = \{(a,b) \in \mathbb{R} \times \mathbb{R} : b \neq 0\}$ forms a group under the operation defined by

$$(a,b)(c,d) = (a+bc, bd)$$

   Show that G is not abelian. Show also that the subsets

$$\begin{aligned} H &= \{(a,b) \in \mathbb{R} \times \mathbb{R} : a = 0\}, \\ K &= \{(a,b) \in G : b > 0\}, \\ L &= \{(a,b) \in G : b = 1\} \end{aligned}$$

   are subgroups of G. Which of these subgroups are abelian?

Definition(A Binary Operation)
A binary operation $*$ on a set G is an operation which, when applied to any elements x and y of the set G, yields an element $x * y$ of G.

Example
For all $x, y \in \mathbb{R}$ (the set of real numbers), we have

$$x + y, x - y \ \text{ and } \ xy \in \mathbb{R}.$$

Thus, $+$, $-$, $\cdot$ are binary operations on the set of real numbers.
However, $x/y \notin \mathbb{R}$ (in general) because the quotient x/y is not defined when $y = 0$. Thus division is not a binary operation on the set of real numbers.

Definition(Commutative Binary Operation)
A binary operation $*$ on a set G is said to be commutative if $x * y = y * x$ for all elements x and y of G.

Example

The operations of $+$ and $\cdot$ on the set $\mathbb{R}$ of real numbers are commutative, since $x+y=y+x$ and $x\times y=y\times x$ for all $x,y\in\mathbb{R}$. However, the operation of subtraction is not commutative, since $x-y\neq y-x$ in general. (Indeed the identity $x-y=y-x$ holds only when $x=y$).

Definition(Associative Binary Operation)

A binary operation $*$ on a set G is said to be associative if $(x*y)*z=x*(y*z)$ for all elements x,y and z of G.

Example

Since

$$(x+y)+z=x+(y+z) \text{ and } (x\times y)\times z=x\times(y\times z)$$

for all real numbers x,y and z.
Thus the operations of $+$ and $\times$ on the set $\mathbb{R}$ of real numbers are associative. However, the operation of subtraction is not associative . For Example, $(1-2)-3=-4$, but $1-(2-3)=2$.

Remark

When a binary operation $*$ is associative it is not necessary to retain the parentheses in expressions such as

$$(x*y)*z \text{ or } x*(y*z).$$

These two expressions may both be written without ambiguity as $x*y*z$.

Definition (Semigroup)

A semigroup consists of a set on which is defined an associative binary operation. We may denote by $(G,*)$ a semigroup consisting of a set G together with an associative binary operation $*$ on G.

Definition (Abelian Semigroup)

A semigroup $(G,*)$ is said to be commutative (or Abelian) if the binary operation $*$ is commutative.

Example

The set of natural numbers $\mathbb{N}$, with the operation of addition, is a commutative semigroup, as is the set of natural numbers $\mathbb{N}$ with the operation of multiplication.

Let $(G, *)$ be a semigroup. Given any element g of G, we define

$$\begin{aligned} g^1 &= g, \\ g^2 &= g * g, \\ g^3 &= g * g^2 = g * (g * g), \\ g^4 &= g * g^3 = g * (g * (g * g)), \\ g^5 &= g * g^4 = g * (g * (g * (g * g))). \end{aligned}$$

In general, we define g^n recursively for all natural numbers n so that $g^1 = g$ and $g^n = g * g^{n-1}$ whenever $n > 1$.

Remark

The value of "a^n" given by the above rule is the nth power of a natural number g. However in the case of the semigroup consisting of the set of natural numbers with the operation of addition, it is not the nth power of a, but is na.

Theorem

Let $(G, *)$ be a semigroup, and let g be an element of G. Then $g^m * g^n = g^{m+n}$ for all natural numbers m and n.

Theorem

Let $(G, *)$ be a semigroup, and let g be an element of G. Then $(g^m)^n = g^{mn}$ for all natural numbers m and n.

Remark

In any semigroup, the value of a product of three or more elements of the semigroup depends in general on the order in which those elements occur in the product (unless the binary operation is commutative), but the value of the product is independent of the manner in which the product is bracketed.

Example

Let $(G, *)$ be a semigroup, and let x, y, z and w be elements of G. We can use the associative property of $*$ to show that the value of a product involving x, y, z, w is independent of the manner in which that product is bracketed.

$$\begin{aligned}(x * (y * z)) * w &= ((x * y) * z) * w \\ &= (x * y) * (z * w) \\ &= x * (y * (z * w)) \\ &= x * ((y * z) * w).\end{aligned}$$

All the above products may therefore be denoted without ambiguity by the expressions $x * y * z * w$ from which the parentheses have been dropped.

Definition (Identity element of a semigroup)

Let $(G, *)$ be a semigroup. An element e of G is said to be an identity element for the binary operation $*$ if

$$e * x = x * e = x \text{ for all elements } x \text{ of } G.$$

Example

The number 1 is an identity element for the operation of multiplication on the set $\mathbb{N}$ of natural numbers.

Example

The number 0 is an identity element for the operation of addition on the set $\mathbb{Z}$ of integers.

Theorem(Uniqueness of Identity in a Set)

A binary operation on a set cannot have more than one identity element.

Proof

Let e and f be identity elements for a binary operation $*$ on a set S. Then

$$e = e * f = f.$$

Thus, there cannot be more than one identity element.

Definition (Monoid)

A monoid consists of a set on which is defined an associative binary operation with an identity element.

It follows that a semigroup is a monoid if and only if it has an identity element.

Definition (Abelian monoid)

A monoid $(G, *)$ is said to be commutative (or Abelian if the binary operation $*$ is commutative.

Example

The set $\mathbb{N}$ of natural numbers with the operation of multiplication is both commutative and associative and the identity element is the natural number 1.

Example

The set $\mathbb{N}$ of natural numbers with the operation of addition is not a monoid, since there is no identity element for the operation of addition that belongs to the set of natural numbers.

Let a be an element of a monoid $(G, *)$. We define $a^0 = e$, where e is the identity element.

Theorem

Let $(G, *)$ be a monoid, and let a be an element of G. Then $a^m * a^n = a^{m+n}$ for all non-negative integers m and n.

Theorem

Let $(G, *)$ be a monoid, and let a be an element of G. Then

$$(a^m)^n = a^{mn}$$

for all nonnegative integers m and n.

Definition (Inverse element of a Monoid)

Let $(G, *)$ be a monoid with identity element e, and let x be an element of G. An element y of G is said to be the inverse of x if $x * y = y * x = e$. An

element x of G is said to be invertible if there exists an element of G which is an inverse of x.

Theorem(Uniqueness of Inverse Element in a Monoid)
An element of a monoid can have at most one inverse.

Proof
Let $(G, *)$ be a monoid with identity element e, and let x, y and z be elements of G. Suppose that

$$x * y = y * x = e \text{ and } x * z = z * x = e.$$

Then

$$\begin{aligned} y = y * e &= y * (x * z) \\ &= (y * x) * z \\ &= e * z \\ &= z \end{aligned}$$

and thus $y = z$.
Thus an element of a monoid cannot have more than one inverse.

Notation
Let $(G, *)$ be a monoid, and let x be an invertible element of G. We shall denote the inverse of x by x^{-1}

Theorem
Let $(G, *)$ be a monoid, and let x and y be invertible elements of G. Then $x * y$ is also invertible, and $(x * y)^{-1} = y^{-1} * x^{-1}$.

Theorem
Let $(G, *)$ be a monoid, let a and b be elements of G, and let x be an invertible element of G. Then $a = b * x$ if and only if $b = a * x^{-1}$. Similarly, $a = x * b$ if and only if $b = x^{-1} * a$.

Note:
Let $(G, *)$ be a monoid, and let a be an invertible element of G. We extend the definition of a^n to negative integers n by defining a^n to be the inverse $(a^q)^{-1}$ of a^q whenever $q > 0$ and $n = -q$.

Theorem
Let $(G, *)$ be a monoid, and let a be an invertible element of G. Then $a^m * a^n = a^{m+n}$ for all integers m and n.

Theorem
Let $(G, *)$ be a monoid, and let a be an invertible element of G. Then $(a^m)^n = a^{mn}$ for all integers m and n.

Definition (Group)
A group consists of a set G together with a binary operation $*$ on G with the following properties:-

(i) $x * (y * z) = (x * y) * z$ for all elements x, y and z of G (i.e., the operation $*$ is associative);

(ii) there exists an element e of G with the property that
$e * x = x * e = x$ for all elements x of G
(i.e., there exists an identity e for the binary operation $*$ on G);

(iii) given any element x of G, there exists an element y of G satisfying

$$x * y = y * x = e$$

(i.e., every element of G is invertible).

We see immediately from this definition that a group can be characterized as a monoid in which every element is invertible.

Definition (Abelian Group)
A group $(G, *)$ is said to be commutative (or Abelian) if the binary operation $*$ is commutative.

Examples

(i) The set of integers with the operation of addition is a commutative group.

(ii) The set of real numbers with the operation of addition is a commutative group.

(iii) The set of non-zero real numbers with the operation of multiplication is a commutative group.

(iv) The set of integers with the operation of multiplication is not a group, since not every element is invertible.
Indeed the only integers that are invertible are $+1$ and -1.

(v) Let n be a natural number, and let $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ with the operation of addition modulo n is a group. However, the set $\mathbb{Z}_n$ with the operation of multiplication modulo n is not a group, since not all elements have multiplicative inverses.

Practice Exercise

Let G and $*$ be the sets and operations defined below respectively. Consider $(G, *)$ and show whether they are groups or not G.

(i) $G = \{\phi, \{a\}, \{b\}, \{a, b\}, * = \cap$ (intersection)

(ii) $G = \{1, 3, 5, 7\}, * = \bigotimes_8$ (multiplication modulo 8)

Definition (Finite Group)

A group G is finite if $|G|$ is finite. The number of elements in a finite group is called its order.

Definition (Subgroup, Proper Subgroup)

A non-empty subset H of a group G is a subgroup of G if H is itself a group with respect to the operation of G. If H is a subgroup of G and $H \neq G$, then H is called a proper subgroup of G. When H is a subgroup of G, we write $H \leq G$.

Equivalently, H is a subgroup of $(G, \cdot)$ if and only if $ab^{-1} \in H$, for all $a, b \in H$ or H is a subgroup of $(G, +)$ if and only if $a - b \in H$ for all $a, b \in H$.

Examples

$H = \{e\}$ and $H = G$ are both subgroups of G called the trivial subgroups of G.

Definition (Cyclic Group)

A group G is cyclic if there is an element $g \in G$ such that for each $t \in G$ there is an integer i with $t = g^i$. Such an element g is called a generator of G.

Examples

(i) $(\mathbb{Z}, +) = \langle 1 \rangle$ is an infinite cyclic group

(ii) $\langle a \rangle = \{a, a^2, a^3, a^4, a^5, a^6 = e\}$ is a finite cyclic group of order 6.

Fact

Let G be a group, and let $g \in G$ be an element of finite order t. Then $|\langle g \rangle|$, the size of the subgroup generated by g, is equal to t.

Fact (Langrange's Theorem)

If G is a finite group and H is a subgroup of G, then $|H|$ divides $|G|$. Hence, if $g \in G$, the order of g divides $|G|$.

Example

All the subgroups of a group of order 6 are of orders 1, 2, 3 or 6.

Fact

Every subgroup of a cyclic group G is also cyclic. In fact, if G is a cyclic group of order n, then for each positive divisor d of n, G contains exactly one subgroup of order d.

Fact

Let G be a group.

(i) If the order of $g \in G$ is t, then the order of g^k is $t/gcd(t, k)$

(ii) If G is a cyclic group of order n and $d \backslash n$, then G has exactly $\phi(d)$ elements of order d. In particular G has $\phi(n)$ generators where gcd denote greatest common divisor, $\phi(d)$ denote the number of integers in the interval $[1, d]$

which are relatively prime to d. The function ϕ is called the Euler phi function (or the Euler totient function).

Practice Exercise

(i) Compute the subgroups $\langle 0\rangle$, $\langle 1\rangle$, $\langle 2\rangle$, $\langle 3\rangle$, $\langle 4\rangle$ and $\langle 5\rangle$ of the group $\mathbb{Z}_6$.

(ii) Which elements are generators for the group $\mathbb{Z}_6$ of part (i).

Definition (Permutation of a set)

Let G be a finite set. Then a permutation of the set G is a bijective mapping from G to itself.

Example

The above definition can be illustrated in form of notations as:

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 2 \\ 1 \\ 4 \\ 3 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \quad \text{or} \quad (1\ \ 2)(3\ \ 4)$$

Theorem

The set of all permutations on a set G is a group under the operation of composition of mappings.

Definition (Symmetric group of n letters)

The group of all permutations on a set G such that $|G| = n$ is called the symmetric group of n letters and is denoted by S_n.

Example:

Let $G = \{1, 2, 3\}$ where $|G| = 3$

Then

$$S_n = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \right\}$$

Remark

We observe that the order of the columns of each permutation in above example is immaterial.
For instance,

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 2 \\ 3 & 2 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 2 & 3 & 1 \\ 1 & 3 & 2 \end{pmatrix}$$

Definition (Binary Operation of Symmetric Group)

If α, β are two elements of S_n such that

$$\alpha = \begin{pmatrix} 1 & 2 & \cdots & n \\ \alpha(1) & \alpha(2) & \cdots & \alpha(n) \end{pmatrix}$$

$$\alpha = \begin{pmatrix} 1 & 2 & \cdots & n \\ \beta(1) & \beta(2) & \cdots & \beta(n) \end{pmatrix}$$

The product (composition) $\beta \cdot \alpha$ is defined as

$$\beta \cdot \alpha = \begin{pmatrix} 1 & 2 & \cdots & n \\ \beta\alpha(1) & \beta\alpha(2) & \cdots & \beta\alpha(n) \end{pmatrix}$$

and

$$\alpha^{-1} = \begin{pmatrix} \alpha(1) & \alpha(2) & \cdots & \alpha(n) \\ 1 & 2 & \cdots & n \end{pmatrix}$$

Example

Consider $\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \beta = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$.

Thus $\beta \circ \alpha = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$

where

$\alpha(1) = 1$, $\alpha(2) = 3$, $\alpha(3) = 2$
$\beta(1) = 2$, $\beta(2) = 1$, $\beta(3) = 3$.

$$\begin{aligned} \beta \circ \alpha(1) &= \beta(1) = 2 \\ \beta \circ \alpha(2) &= \beta(3) = 3 \\ \beta \circ \alpha(3) &= \beta(2) = 1 \end{aligned}$$

and

$$\alpha^{-1} = \begin{pmatrix} 1 & 3 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

where $\alpha(1) = 1$, $\alpha(2) = 3$, $\alpha(3) = 2$.
Similarly,

$$\begin{aligned} \alpha \circ \beta &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \end{aligned}$$

We see that $\alpha \circ \beta \neq \beta \circ \alpha$.

Thus, S_3 is a noncommutative group. The identity element of S_3 is $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$.

Definition (Cyclic Permutation)

Let $x_1, x_2, \ldots, x_n$ be distinct symbols.
Consider the permutation

$$\begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_{n-1} & x_n \\ x_2 & x_3 & x_4 & \cdots & x_n & x_1 \end{pmatrix}$$

is called a cyclic permutation.

Example

$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$ is a cyclic permutation of S_4.

Notice that $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix}$ cycle of length 4.

Practice Exercise

Let $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$ and $\beta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}$.
Compute $\alpha^{-1}\beta$, $\beta^{-1}\alpha$, $\alpha\beta^{-1}$ and $\beta\alpha^{-1}$.

Theorem

Any permutation in S_n can be written as a product of disjoint cycles.

Example

In S_7, $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 6 & 1 & 3 & 7 & 2 & 5 \end{pmatrix} = (1\ 4\ 3)(2\ 6)(5\ 7)$.

Definition (Transposition)

A cycle of length 2 is called a transposition i.e., the permutation in S_n which interchanges the number i, j and leaves the other elements fixed.

Example

In S_6, $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 1 & 4 & 5 & 6 \end{pmatrix} = (1\ 3)$ is a cycle of length 2.

Remark

From the above theorem we can conclude that a cycle of length n can be written as a product of $(n-1)$ transpositions i.e.,

$$(a_1, a_2 \cdots a_n) = (a_1a_2)(a_1a_3)(a_1a_4)\cdots(a_1a_n)$$

Example

In S_8,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 3 & 4 & 5 & 6 & 7 & 8 & 1 \end{pmatrix} = (1\ 2\ 3\ 4\ 5\ 6\ 7\ 8)$$

is a cycle of length 8 can be broken up into product of transpositions as

$$(1\ 8)(1\ 7)(1\ 6)(1\ 5)(1\ 4)(1\ 3)(1\ 2).$$

Generally, any permutation in S_n can be written as a product of transpositions.

Theorem (Cayley)

Any group $(G, *)$ of finite order n is isomorphic to a subgroup of the symmetric group of degree n.

Definition (Homomorphism)

Let $(G_1, *)$ and $(G_2, *)$ be semigroups, monoids or groups. A function $f : G_1 \longrightarrow G_2$ from G_1 to G_2 is said to be a homomorphism if $f(x * y) = f(x) * f(y)$ for all elements x and y of G_1.

Example

Let r be an integer, and let $f : \mathbb{Z} \to \mathbb{Z}$ be a function from the set of integers to itself defined by $f(n) = rn$ for all integers n. Then f is a homomorphism from the group $(\mathbb{Z}, +)$ to itself, since

$$f(m + n) = r(m + n) = rm + rn = f(m) + f(n)$$

for all integers m and n.

Example

Let $\mathbb{R}^*$ denote the set of non-zero real numbers, let a be a non-zero real number, and let $f : \mathbb{Z} \to \mathbb{R}^*$ be the function defined by $f(n) = a^n$ for all integers n. Then $f : \mathbb{Z} \to \mathbb{R}^*$ is a homomorphism from the group $(\mathbb{Z}, +)$ of integers under addition to the group $(\mathbb{R}^*, \times)$ of non-zero real numbers under multiplication, since

$$f(m + n) = a^{m+n} = a^m a^n = f(m)f(n)$$

for all integers m and n.

Definition (Isomorphism)

Let $(G_1, *)$ and $(G_2, *)$ be semigroups, monoids, or groups. A function $f : G_1 \to G_2$ from G_1 to G_2 is said to be an isomorphism if it is both a homomorphism and a bijective function.

Theorem

Let $(G_1, *)$ and $(G_2, *)$ be semigroups, monoids or groups. Then the inverse $f^{-1} : G_2 \to G_1$ of any isomorphism $f : G_1 \to G_2$ is itself an isomorphism.

Definition (Isomorphic)

Let $(G_1, *)$ and $(G_2, *)$ be semigroups, monoids or groups. If there exists an isomorphism from $(G_1, *)$ to $(G_2, *)$ then $(G_1, *)$ and $(G_2, *)$ are said to be isomorphic.

Summary

Algebraic structures such as groupoids, semigroups, monoids and groups are considered as well as abelian or commutative groups. Properties of groups are given and examples from numbers, sets, residue classes, symmetric groups and cyclic groups are studied.

Conditions for a subset of a group to be a subgroup are given and Lagrange's theorem is used to determine subgroups of given finite groups.

Mappings between groups which presence the binary operations of the groups are called homoorphisms. Homomorphisms which are one-to-one or onto or both are studied.

Post-Test

See Pre-Test at the beginning of the lecture.

Supplementary Reading

1. M. Artin, Algebra, Groups, (Chapter 2), Prentice-Hall, Upper Saddle River, New Jersey 074548, 1991, pp. 38-61.

2. A.O. Kuku, Abstract Algebra, Ibadan University Press, 1980, Chapter 3, pp. 95-125.

3. S.A. Ilori and O. Akinyele, Elementary Abstract and Linear Algebra, Ibadan University Press, 1986, pp 157-192.

4. F. Ayres, Modern Algebra, Schaum's Outline Series.

5. G. Birkoff and S. MaClane, A survey of Modern Algebra, Macmillan Co. 1965.

6. J.B. Fraleigh, A First Course in Abstract Algebra, Addison Wesley 1968.

7. E.H. Connell, Elements of Abstract and Linear Algebra. (http://www/math.miami.edu/∼ec/book/).

LECTURE 14

Boolean Algebra

Introduction

We shall study a special class of relations called partially ordered sets and a class of algebras. Finally, we shall mention an application in switching circuits.

Objectives

At the end of this lecture you should be able to do the following:
- identify partially ordered sets (posets),
- identify the lower and upper bounds, greatest lower and least upper bounds of subsets of posets,
- give the definition, examples and elementary properties of lattice and Boolean algebras, and
- construct an application in switching circuit or computer theory.

Pre-Test

1. Let $S = \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$ and let $\sim$ be a relation defined on S by $n \sim m$ if m is a multiple of n.
   Is $(S, \sim)$ a partially ordered sets?

2. Let $(\mathbb{R}, \leq)$ be a poset where $\mathbb{R}$ is the set of real numbers. Consider the open interval (a, b) as a subset of $\mathbb{R}$. Find,

   (i) the set of lower bounds of (a, b);

   (ii) the set of greatest lower bounds of (a, b);

   (iii) the set of upper bounds of (a, b);

   (iv) the set of least upper bounds of (a, b).

3. Let $(\mathbb{R}, \leq)$ be a poset where $\mathbb{R}$ is the set of real numbers. Consider the closed interval $[a, b]$ as a subset of $\mathbb{R}$. Find for the subset $[a, b]$ the set of (i) lower bounds, (ii) greatest lower bounds, (iii) upper bounds,

   (iv) least upper bounds.

4. Complete the following operation tables in a Boolean algebra.

$\vee$	0	1	a	a'
0	0	1	a	a'
1	1			
a	a			
a'	a'			

$\wedge$	0	1	a	a'
0		0		
1	0	1	a	a'
a		a		
a'		a'		

5. Show that 1 is unique in a Boolean algebra.

6. Assume that the tables below show Boolean functions:

$+$	i	j	k	h
i	i	j	k	h
j	j	j	j	j
k	k	j	k	j
h	h	j	j	h

$\bullet$	i	j	k	h
i	i	i	i	i
j	i	j	k	h
k	i	k	k	i
h	i	h	i	h

   Identify 0 and 1 and find out which laws are presented in the tables.

7. Interpret Fig 14.1 in symbols; then simplify and re-draw the circuit.

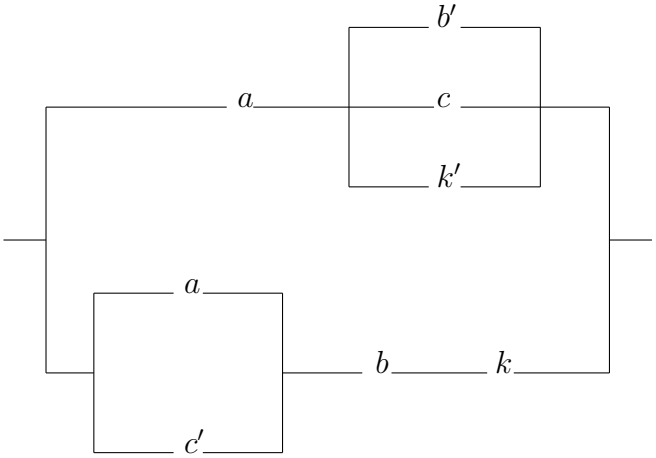


Fig. 14.1

8. Interpret and simplify the circuit of Fig. 14.2

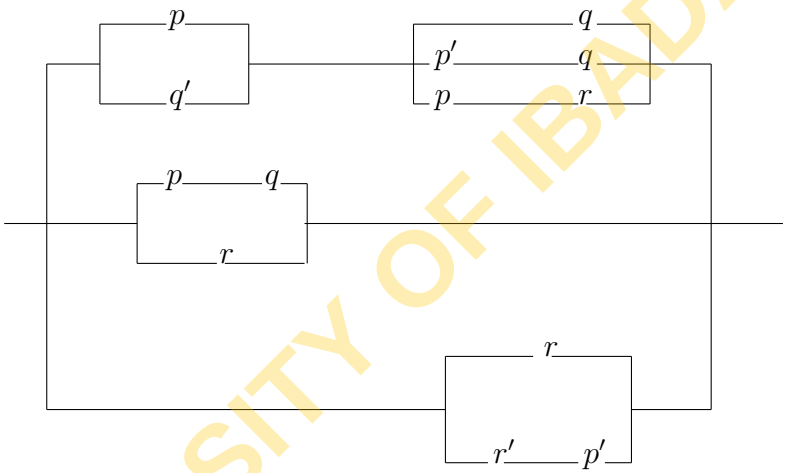


Fig. 14.2

9. Invent some examples of your own.

Partial Orders and Lattices

Definition (Anti-symmetric binary relation)

Let A be a set. A binary relation R on A is said to be anti-symmetric if it has the following property:

If x and y are elements of A; and if xRy and yRx, then $x = y$.

Definition (Partial Order)

A partial order on a set is a relation on that set which is reflexive, transitive and anti-symmetric.

Let $\leq$ denote a relation on a set A. We see that this relation is a partial order on a set if and only if it has the following three properties:

(i) $x \leq x$ for all elements x of A,

(ii) if x, y, and z are elements of A, and if $x \leq y$ and $y \leq z$; then $x \leq z$;

(iii) if x and y are elements of A, and if $x \leq y$ and $y \leq x$; then $x = y$.

Example

The relation $\leq$ ("less than or equal to") is a partial order on the set $\mathbb{R}$ of real numbers. (It clearly possesses all three properties listed above). It is

also a partial order when considered as a relation on the set $\mathbb{Z}$ of integers, or on the set $\mathbb{N}$ of natural numbers.

Example

Let A be a set. The relation $\subset$ is a partial order on the power set $\mathcal{P}(A)$ of A, where subsets B and C satisfy $B \subset C$ if and only if B is a subset of C (i.e. if and only if every element of B belongs also to C).

Definitions

A partially ordered set (or poset) $(A, \leq)$ consists of a set A, which is provided with a partial order $\leq$ defined on the set.

Let $(A, \leq)$ be a partially ordered set, and let B be a subset of A. An element l of A is said to be a lower bound of B if $l \leq b$ for all elements b of B.

An element l of A is said to be the greatest lower bound of B if l is a lower bound of B and if $l' \leq l$ for all lower bounds l' of$B$. If such a greatest lower

bound exists, we shall denote it by $glb(B)$. It is worth noting that a subset B of A can have at most one greatest lower bound.

We can define in a similar fashion the notion of a least upper bound of a subset of A. An element u of A is said to be an upper bound of a subset B of A if $b \leq u$ for all elements b of B. An element u of A is said to be the least upper bound of B if u is an upper bound of B and if $u \leq u'$ for all upper bounds u' of B. A subset B of A can have at most one least upper bound. If such a least upper bound exists, we shall denote it by $lub(B)$.

Example

Consider the partially ordered set $(\mathbb{R}, \leq)$. Any finite subset B of $\mathbb{R}$ has both a greatest lower bound and a least upper bound. The greatest lower bound of B in this case is the smallest real number belonging to B, and the least upper bound is the largest real number belonging to B.

Example

Let A be a set, and let $\mathcal{P}(A)$ be the power set of A (i.e., the set whose elements are the subsets of A). Then $(\mathcal{P}(A), \subset)$ is a partially ordered set, the relation $\subset$ is a partial order on the power set $\mathcal{P}(A)$ of A. Given subsets B and C of A one can readily verify that

$$glb\{B, C\} = B \cap C,\ lub\{B, C\} = B \cup C.$$

Indeed, $B \cap C \subset B$ and $B \cap C \subset C$, and therefore $B \cap C$ is a lower bound of $\{B, C\}$. Moreover, if D is any subset of A that is a lower bound of $\{B, C\}$ then $D \subset B$ and $D \subset C$, hence the elements of D must belong to both B and C, hence $D \subset B \cap C$. This shows that $glb\{B, C\} = B \cap C$. A similar argument shows that $lub\{B, C\} = B \cup C$.

Example

Let $(\mathbb{N}, \leq)$ be the partially ordered set (poset) consisting of the set $\mathbb{N}$ of natural numbers, together with the usual partial order $\leq$. Let B be the subset of $\mathbb{N}$ consisting of all the even natural numbers (i.e., $B = \{2, 4, 6, 8, \dots\}$). The set B has a greatest lower bound. Indeed $glb(B) = 2$. But the set B has no least upper bound: Indeed the set has no upper bound: no natural number

has the property that it is greater than or equal to all even natural numbers.

Definition (Lattice)

A partially ordered set $(A, \leq)$ is said to be a lattice if, given any two elements x and y of the set A, there exists an element $glb\{x, y\}$ of A that is, the greatest lower bound of the set $\{x, y\}$ and an element $lub\{x, y\}$ that is, the least upper bound of the set $\{x, y\}$.

Example

$(\mathbb{R}, \leq)$, $(\mathbb{Z}, \leq)$, $(\mathbb{N}, \leq)$ are lattices (where two numbers x and y satisfy $x \leq y$ if and only if x is less than or equal to y).

Example

Let A be a set, and let $\mathcal{P}(A)$ denote the power set of A. Then $(\mathcal{P}(A), \subset)$ is a lattice.
Indeed, if B and C are elements of $\mathcal{P}(A)$ then they are subsets of A. Moreover, we have already seen that $glb\{B, C\} = B \cap C$, $lub\{B, C\} = B \cup C$, and $B \cap C$ and $B \cup C$ are elements of $\mathcal{P}(A)$ (since they are obviously subsets of A). It follows that $(\mathcal{P}(A), \subset)$ contains the greatest lower bound and least upper bound of the set $\{B, C\}$ for all elements B and C of $\mathcal{P}(A)$ (i.e. for all subsets B and C of A).

Practice Exercise

Let $S(V)$ be the set of all subspaces of a vector space V and partially ordered by set inclusion. Show how this leads to the lattice of subspaces.

Definition (Distributive lattice)

A lattice is distributive if

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$$

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$$

where

$$a \wedge b = glb\ \{a, b\} \text{ and } a \vee b = lub\ \{a, b\}.$$

Definition (Complemented lattice)

A lattice is complemented if it contains distinct elements 0 and 1 such that $0 \leq a \leq 1$ for every element a in the lattice, and if each element a has a complement a' with the property that $a \wedge a' = 0$ and $a \vee a' = 1$.

Definition (Boolean algebra)

A Boolean algebra is a complemented distributive lattice.

An Application

The application likely to be of greatest interest is that of electrical or electronic switching devices. A switch is either on or off, giving the 1 and 0. An inversion device like a break-contact ensures the compliments: if $x = 0$, then $x' = 1$, or $x = 1$ gives $x' = 0$. This means that if a switch is marked x', it is on when x is off or vice-versa.

Two (or more) switches in series give $x \cdot y$ (both)

Fig. 14.3

Two (or more) switches in parallel give $x + y$ (either)

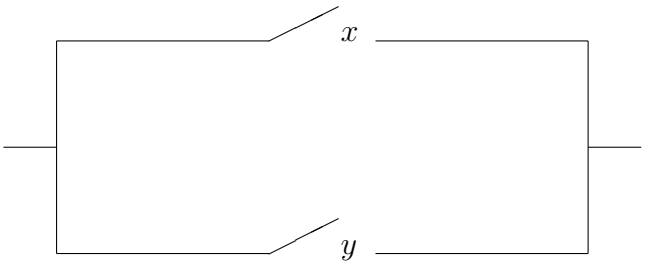


Fig. 14.4

The manipulation of the hardware for constructing circuits is obviously not within the province of abstract algebra, but the algebra can be a real help in finding the simplest circuit to carry specified instructions. ("Simplest" is usually equated to the smallest possible number of switches). Example:

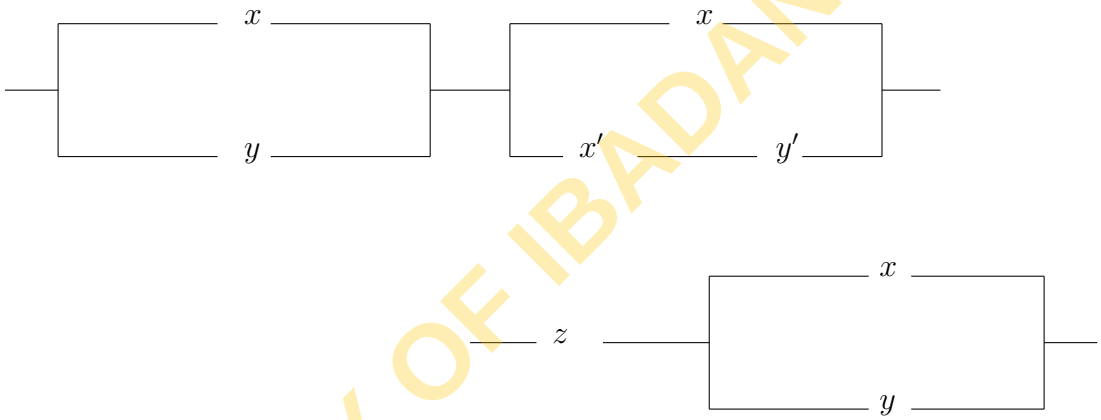


Fig. 14.5

The two diagrams above represent circuit with the same effect. The formulae are $z(x+y)$ and $(x+y)(z+x'y')$.
To prove, analytically, that they are equivalent:

Proof:

$(x+y)(z+x'y')$	Reasons (laws used)
$=(x+y)z+(x+y)x'y'$	first distributive
$=(x+y)z+x'y'(x+y)$	commutative
$=(x+y)z+x'xy'+x'y'y$	first distributive and commutative
$=z(x+y)+xx'y'+x'yy'$	commutative, three times
$=z(x+y)+0+0$	complementation twice
$=z(x+y)$	identity +

Practice Exercise

1. Design a circuit equivalent to the formula $(xy+y'z)(x+y)$ and show that it is replaceable by $x(y+z)$; represent both by diagrams.
2. Show that $xy+x'z+yz=xy+x'z=(x+z)(x'+y)(y+z)$ design all three circuits.

Summary

We first consider partially ordered sets (posets) which satisfy the reflexive, anti-symmetric and transitive laws. We then studied the following special elements of a poset.
(i) lower bounds and upper bounds
(ii) greatest lower bounds and least upper bounds
Properties of a Boolean algebra and its applications are then studied.

Post-Test

See Pre-Test at the beginning of the lecture.

Supplementary Reading

1. A.O. Kuku, Abstract Algebra, Ibadan University Press, 1980, pp. 53-61.
2. S.A. Ilori and O. Akinyele, Elementary Abstract and Linear Algebra, Ibadan University Press, 1986, pp 138-141, 151-152, 155-156.
3. S. Lipschutz, Set Theory and Related Topics, Schaum's Outline Series.
4. F. Ayres, Modern Algebra, Schaum's Outline Series.
5. S. Lipschutz, General Topology, Schaum's Outline Series Chapter 3.
6. G. Birkoff and S. MaClane, A survey of Modern Algebra, Macmillan Co. 1965.
7. J.B. Fraleigh, A First Course in Abstract Algebra, Addison Wesley 1968.
8. G. Takenti and W.M. Zaring, Axiomatic Set Theory, Boolean Algebra (Chapter 1), Springer-Verlag, New York Inc.
9. E.H. Connell, Elements of Abstract and Linear Algebra. (http://www/math.miami.edu/∼ec/book/).

LECTURE 15

Rings and Fields

Introduction

Unless it is well understood, you should revise "groups" before attempting this lecture. The point was made, in lecture 13, that a group has only one operation, and may be called "one process" algebra. Two operations were used simultaneously in Boolean algebra, and two operations are also required by rings and fields. This lecture therefore qualify as "two-process" algebras. The law which links the two operations is the distributive law.

Objectives
At the end of this lecture you should be able to:
• itemize the requirements of an additive group,
• know what must be true before ring structure is established,
• itemize the requirements that are necessary for a field, but not essential for a ring, and
• identify some unusual features of Boolean algebra that prevent it qualifying as a field.

Pre-Test

See Post-Test at the end of the lecture.

Here we considered with two operations, called addition and "multiplication", that can both be used on a given set S of at least two elements. The rule by which the pair of operations is linked is called the distributive law. To allow for non-commutative multiplication, it takes two forms:

(1) $p * (q \oplus r) = (p * q) \oplus (p * r)$

(2) $(q \oplus r) * p = (q * p) \oplus (r * p)$ for all $p, q, r \in S$.

Both must be true, but if $*$ is commutative, the results are not different.
You have already met this law in studying sets and Boolean algebra. If you have forgotten, test for the numbers 5,6 and 8 with ordinary addition and multiplication. To indicate the wide application of these laws, general signs $\oplus$ and $*$ are used instead of $+$ and $\times$.

Practice Exercise

- If $*$ means $\div$ (division) and $\oplus$ means $+$, are either of the distributive laws, or both, true for $p, q, r \in \mathbb{Z}$?
- Are the laws both true, and do they give the same results if $*$ means matrix multiplication, $\oplus$ means matrix addition, and

$$p = \begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix}, \; q = \begin{pmatrix} 1 & -2 \\ 3 & -1 \end{pmatrix}, \; r = \begin{pmatrix} 3 & 1 \\ -1 & 2 \end{pmatrix}?$$

Note: You are advised not to read on until you have made yourself completely familiar with the distributive law. Also recall that it is not new: you have always used it for factorisation and for expansion of brackets in ordinary algebra.

Ring

The minimum requirements for a ring structure are:

(1) A set of at least two elements with defined operation $\oplus$ must form a commutative group (or additive group), i.e., there is closure, associativity, a zero identity element, and unique inverses for every element.

(2) The same elements and defined operation $*$ must give closure and associativity.

(3) Both distribution laws must hold.

Notice that inverses are not required for $*$, and that the zero element cannot possibly have an inverse for $*$; also notice that divisors of zero are not forbidden.

Rings are of several kinds. A ring may, for $*$, have an identity element e, usually now called the unity element (to distinguish it from the zero element).

If e exists the ring is a ring with unity; e will be its own inverse, and there may be other elements that have inverses. If one of the element p has an inverse p^{-1} such that $p * p^{-1} = p^{-1} * p = e$, then p is called a regular element or a "unit" of the ring and p^{-1} must also be a unit of the ring.

The operation $*$ may be commutative, when the ring is a commutative ring. If there is both a unity element and commutativity for $*$ we have a commutative ring with unity. This is sometimes called an integral domain because the integers, $\mathbb{Z}$, give an example of this structure. In fact, an integral domain is a commutative ring with unity and with no proper zero divisors.

Examples

Examples of rings are found in both infinite and finite sets.
Thus, for the infinite set of all 2×2 matrices with $\oplus$ and $*$ meaning matrix addition and multiplication:

(1) That $\oplus$ is a commutative group, you should verify this through lecture 13.

(2) With $*$ there is closure and associativity.

(3) Do the distributive laws hold for matrices of real numbers.

$$p = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad q = \begin{pmatrix} e & f \\ g & h \end{pmatrix}, \quad r = \begin{pmatrix} i & j \\ k & l \end{pmatrix}?$$

We show the working for the first and ask you to check the second law in the same way,

$$\begin{aligned} p * (q \oplus r) &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e+i & f+j \\ g+k & h+l \end{pmatrix} \\ &= \begin{pmatrix} ae+ai+bg+bk & af+aj+bh+bl \\ ce+ci+dg+dk & cf+cj+dh+dl \end{pmatrix} \qquad (1) \\ (p*q) \oplus (p*r) &= \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix} \oplus \begin{pmatrix} ai+bk & aj+bl \\ ci+dk & cj+dl \end{pmatrix} \\ &= (1) \quad \text{by inspection.} \end{aligned}$$

So, $p * (q \oplus r) = (p * q) \oplus (p * r)$.

(4) There is, moreover, a unity element $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.
Thus the set of 2×2 matrices is a ring with a unity.

(5) The set $\mathbb{Z}_n$ with addition and multiplication performed modulo n is a finite commutative ring with unity. But it is not an integral domain in general, unless n is prime.
For example, $(\mathbb{Z}_6, \oplus, *)$ is not an integral domain.
Since it has proper zero divisors, $\bar{2} \cdot \bar{3} = \bar{6} = \bar{0}$, $\bar{2} \neq 0$, $\bar{3} \neq 0$, whereas $(\mathbb{Z}_7, \oplus, *)$ is an integral domain because it is commutative ring with unity but with no proper zero divisors.

Subrings

Subrings of the ring in (4) above are found by selecting $p = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ etc. for special conditions.
For example,

(i) if a, b, c, d are even integers, there will be no unity element and we have a subring without unity.

(ii) For $a, b, c, d \in \mathbb{Z}$ it is possible to have a subring that is commutative and has a unity element; but regular elements (units) of the ring, other than I, can exist only for those elements p for which the determinant $\triangle = 1$ or -1. Why?

An example of a finite ring is $\mathbb{Z}_6$. You should satisfy yourself that this is a commutative ring with unity. Are there any regular elements?
Hint: 2, 3 and 4 are divisors of zero.

Practice Exercise

- When we work in a ring, unusual things can happen. Show that for $\mathbb{Z}_6$ the equation $x^2 + 3x + 2 = 0$ has 4 solutions.
Hint: test by substituting for x each of the values 0, 1, 2, 3, 4, 5 in turn.
Other rings to be tested are:

- The integers $\mathbb{Z}$ with $+$ and $\times$. Which are the regular elements?

- Even integers with $+$, and $*$ defined as $a * b = \frac{1}{2}$(product ab).

Field
The requirements for a field structure are:

(1) A ring with a unity.

(2) For operation $*$, every element p of the set, except zero, must have a unique inverse p^{-1} also in the set. This implies that there cannot be divisors of zero.
Usually $*$ is commutative, but if not, the field is called a skew field (or division ring).

Practice Exercise

- Do non-singular 2×2 matrices of numbers in $\mathbb{R}$ form a skew field? Test in detail.

Example
Complex numbers have the form $a + ib$ such that $a, b \in \mathbb{R}$ and $i^2 = -1$.
Do complex numbers form a field with $+$ and $\times$ defined as below?
(Equality: $a + ib = c + id$ means $a = c$ and $b = d$).

With addition

(i) $(a + ib) + (c + id) = a + c + i(b + d)$, which is also a complex number, so there is closure.

(ii) $a = b = 0$ gives the zero element 0.

(iii) $a + ib + (-a - ib) = 0$, so there are unique inverses for every $a + ib$.

(iv) $+$ is both associative and commutative since order of addition of real numbers is indifferent.

Hence complex numbers with $+$ are an Abelian group.

With multiplication

(i) $(a+ib)(c+id) = ac - bd + i(ad+bc)$, which is also a complex number, so there is closure.

(ii) $a = 1$ and $b = 0$ gives the unity element 1.

(iii) $(a+ib)(c+id) = 1$ if $ac - bd = 1$ and $ad + bc = 0$.
Check that this gives

$$c = \frac{a}{a^2+b^2}, \quad d = \frac{-b}{a^2+b^2}.$$

Unique inverses are thus possible for all elements except for the case $a = b = 0$ (the zero element).

(iv) Testing associativity:

$$\begin{aligned}
& \{(a+ib)(c+id)\}(e+if) \\
= \; & \{ac - bd + i(ad+bc)\}(e+if) \\
= \; & ace - bde - adf - bcf + i(acf - bdf + ade + bce). \\
& (a+ib)\{(c+id)(e+if)\} \\
= \; & (a+ib)\{ce - df + i(cf+de)\} \\
= \; & ace - adf - bcf - bde + i(bce - bdf + acf + ade).
\end{aligned}$$

Yes, there is associativity.

(v) Testing Commutativity

$$\begin{aligned}
& (a+ib)(c+id) = (ac-bd) + i(ad+bc) \\
& (c+id)(a+ib) \\
= \; & ca - db + i(da+cb) \\
= \; & (ac-bd) + i(ad+bc)
\end{aligned}$$

Yes, there is commutativity.

(vi) Testing one distributive law (the other will be the same):

$$\begin{aligned} & (a+ib)\{(c+id)+(e+if)\} \\ = \; & (a+ib)\{(c+e)+i(d+f)\} \\ = \; & ac+ae-bd-bf+i(bc+be+ad+af). \\ & (a+ib)(c+id)+(a+ib)(e+if) \\ = \; & ac-bd+i(ad+bc)+ae-bf+i(af+be) \\ = \; & ac-ae-bd-bf+i(bc+be+ad+af) \end{aligned}$$

Yes, there is distributivity.

Conclusion: all rules are obeyed, so we have a field.

Summary
We itemized the requirements of an additive group, a ring and field, and gave some examples to demonstrate necessary requirements of a field which are not essentially for a ring. Lastly, we identified some strange features of Boolean algebra which disqualified it from being a field.

Post-Test

Find out whether the following examples constitute fields or rings and describe carefully.
Choose $\oplus$ and $*$ suitably.

(a) Rational numbers $\mathbb{Q}$

(b) $\mathbb{Z}_3$, integers modulo 3.

(c) Real numbers $\mathbb{R}$

(d) $\mathbb{Z}_4$, integers modulo 4

(e) Integers $\mathbb{Z}$

(f) $\mathbb{Z}_5$, integers modulo 5

(g) Natural numbers $\mathbb{N}$.

(h) $\mathbb{Z}_7$, integers modulo 7.

(i) Infinite set of elements of the form $p + q\sqrt{2}$ for

(1) p, q in $\mathbb{Z}$

(2) p, q in $\mathbb{Q}$

(j) $\{0, 2, 4, 6, 8\}$ modulo 10

(k) Matrices $\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$ for $k = 0, 1, -1$.

Supplementary Reading

1. R. Horn and C. Johnson, Matrix Analysis, Cambridge University Press, 1985.

2. N. Jacobson, Basic Algebra I, Second Edition, W.H. Freeman, 1985.

3. R.M. Fyfe and D. Woodrow, Basic Ideas of Abstract Mathematics, University of London Press, 1969. SBN 34007972X.

4. A.O. Kuku, Abstract Algebra, Ibadan University Press, 1980, pp. 53-61.

5. S.A. Ilori and O. Akinyele, Elementary Abstract and Linear Algebra, Ibadan University Press, 1986.

6. J.B. Fraleigh, A First Course in Abstract Algebra, Addison Wesley Publishing Company Inc, Reading, Massachusetts.

7. E.H. Connell, Elements of Abstract and Linear Algebra. (http://www/math.miami.edu/∼ec/book/).