Linear Algebra Notes

by

James L. Moseley

SPACE: THE FINAL FRONTIER

A SERIES OF CLASS NOTES TO INTRODUCE LINEAR AND NONLINEAR PROBLEMS
TO ENGINEERS, SCIENTISTS, AND APPLIED MATHEMATICIANS

LINEAR CLASS NOTES:

A COLLECTION OF HANDOUTS

FOR

REVIEW AND PREVIEW

OF LINEAR THEORY

INCLUDING FUNDAMENTALS

OF

LINEAR ALGEBRA

by

James L. Moseley

CONTENTS

A SERIES OF CLASS NOTES TO INTRODUCE LINEAR AND NONLINEAR PROBLEMS TO ENGINEERS, SCIENTISTS, AND APPLIED MATHEMATICIANS

LINEAR CLASS NOTES:
A COLLECTION OF HANDOUTS FOR REVIEW AND PREVIEW
OF LINEAR THEORY
INCLUDING FUNDAMENTALS OF
LINEAR ALGEBRA

CHAPTER 0

Introductory Material

1. Teaching Objectives for the Linear Algebra Portion of the Course

2. Sample Lectures for Linear Algebra

TEACHING OBJECTIVES
for the
LINEAR ALGEBRA PORTION
of Math 251 (2-4 weeks)

1. Help engineering students to better understand how mathematicians do mathematics, including understanding the concepts of concrete and abstract spaces.
2. Know how to do matrix operations with real and complex matrices. This may require some remedial work in complex arithmetic including the conjugate of a complex number.
3. Require that students memorize the definition of the abstract concept of a Vector (or Linear) Space as an abstract algebraic structure and that of a subspace of a vector space. The definition of a vector space includes the eight axioms for Vector Space Theory. The general definition of a Subspace in mathematics is a set with the same structure as the space. For a Vector Space, this requires closure of vector addition and multiplication by a scalar.
4. Be able to solve $A\vec{x} = \vec{b}$ for all three cases (unique solution, no solution, and infinte number of solutions). Understand what a problem is in mathematics, especially scalar and vector equations.
5. Develop some understanding of the definitions of Linear Operator, Span, Linear Independence, Basis, and Dimension. This need not be indepth, but exposure to all of the definitions is important. Understanding begins with exposure to the definitions. This orients students toward $A\vec{x} = \vec{b}$ as a mapping problem and understanding dimension as an algebraic concept, helping them not to get stuck in 3-space.
6. Know how to compute determinants by both using the Laplace expansion and by using Gauss elimination.
7. Know how to compute the inverse of a matrix. A formula for a 2x2 can be derived.
8. Know the properties of a nonsingular (square) matrix and other appropriate theory (without proofs).

SAMPLE LECTURES FOR LINEAR ALGEBRA

We begin with an algebraic rather than a geometric approach to vectors. This will help the engineers not to get stuck in 3-space. If you start with geometry and then start talking about n-space, the engineers decide that what you are saying has nothing to do with engineering and quit listening. "Help I'm stuck in three space and I can't get out."

DO NOT PRESENT IN CLASS ANY OF THE MATERIAL IN CHAPTER 1. Have them read it on their own. This material provides background to help the students to convert from a geometric approach to an algebraic approach to vectors. This makes it easier to move on to 4,5,...n... dimensions but is not necessary to go over in class.

Lecture #1 "SPACE, THE FINAL FRONTIER"

Handout#1 Page 1 and 2 (on one sheet of paper) of any old exam.

To a mathematician, a **space** is a set plus structure. There are two kinds: **Concrete** and **Abstract**. The number systems, $\mathbf{N}\subseteq\mathbf{Z}\subseteq\mathbf{Q}\subseteq\mathbf{R}\subseteq\mathbf{C}$, are sets with structure and hence are spaces. However, we do not call them spaces, we call them number systems. We may begin with the Peano Postulates, and then construct successively the positive integers (natural numbers) **N**, the integers **Z**,the rational numbers **Q**, and the reals **R** and the complexes **C**. However, in a real analysis course, we follow Euclid and list the axioms that uniquely determine **R**. These fall into three categories: the field (algebraic) axioms, the order axioms, and the least upper bound axiom (LUB). These are sufficient to determine every real number so that **R** is concrete (as are all the number systems). If we instead consider only the field axioms, we have the axioms for the abstract algebraic structure we call a **field**. We can define a field informally as a number system where we can add, subtract, multiply, and divide (except by zero of course). (Do not confuse an algebraic field with a scalar or vector field that we will encounter later in Vector Calculus.) Since an algebraic field is an abstract space, we give three examples: **Q**, **R**, and **C**. **N** and **Z** are not fields. Why?

A **matrix** is an array of elements in a field. Teach them to do matrix algebra with **C** as well as **R** as entries. Begin with the examples on the old exam on page 1. That is, teach them how to compute all of the matrix operations on the old exam. You may have to first teach them how to do complex arithmetic including what the **conjugate** of a complex number is. Some will have had it in high school, some may not. The True/False questions on page 2 of the old exam provide properties of Matrix Algebra. You need not worry about proofs, but for false statements such as "Matrix multiplication of square matrices is commutative" you can provide or have them provide counter examples. DO NOT PRESENT THE MATERIAL IN CHAPTER 2 IN CLASS. It provides the proofs of the matrix algebra properties.

Lecture #2 VECTOR (OR LINEAR) SPACES

Handout#2 One sheet of paper. On one side is the definition of a vector space from the notes. On the other side is the definition of a subspace from the notes.

Begin by reminding them that a space is a set plus structure. Then read them the definition of a vector space. Tell them that they must memorize this definition. This is our second example of an **abstract algebraic space**. Then continue through the notes giving them examples of a vector space including $\mathbf{R}^n$, $\mathbf{C}^n$, matrices, and function spaces.

Then read them the definition of a subspace. Tell them that they must memorize this definition. Stress that a subspace is a vector space in its own right. Why? Use the notes as appropriate, including examples (maybe) of the null space of a Linear Operator.

Lecture #3&4 SOLUTION OF $A\vec{x} = \vec{b}$

Teach them how to do Gauss elimination. See notes. Start with the real 3x3 Example#1 on page 8 of Chapter 4 in the notes. This is an example with exactly one solution. Then do the example on page 15 with an arbitrary $\vec{b}$. Do both of the $\vec{b}$'s. One has no solution; one has an infinite number of solutions. Note that for a 4x4 or larger, geometry is no help. They will get geometry when you go back to Stewart.

Lecture #5&6 REST OF VECTOR SPACE CONCEPTS

Go over the definitions of Linear Operators, Span, Linear Independence, Basis Sets, and Dimension. You may also give theorems, but no proofs (no time). Two examples of Linear operators are $T:\mathbf{R}^n \to \mathbf{R}^m$ defined by $T(\vec{x}) = A\vec{x}$ and the derivative $D:\mathcal{A}(\mathbf{R},\mathbf{R})\to\mathcal{A}(\mathbf{R},\mathbf{R})$ where if $f\in\mathcal{A}(\mathbf{R},\mathbf{R}) = \{f\in\mathcal{F}(\mathbf{R},\mathbf{R})$: f is analytic on $\mathbf{R}$) and $\mathcal{F}(\mathbf{R},\mathbf{R})$ is the set of all real valued functions of a real variable, and we define D by $D(f) = \frac{df}{dx}$.

Lecture #7 DETERMINANTS AND CRAMER'S RULE

Teach them how to compute determinants and how to solve $A\vec{x} = \vec{b}$ using Cramer's Rule.

Lecture #8 INVERSES OF MATRICES

Teach them how to compute (multiplicative) inverses of matrices. You can do any size matrix by augmenting it with the (multiplicative) identity matrix I and using Gauss elimination. I do an arbitrary 2x2 to develop a formula for the inverse of a 2x2 which is easy to memorize.

This indicates two weeks of Lectures, but I usually take longer and this is ok. Try to get across the idea of the difference in a concrete and an abstract space. All engineers need Linear Algebra. Some engineers need Vector Calculus. If we have to short change someplace, it is better the Vector Calculus than the Linear Algebra. When you go back to Stewart in Chapter 10 and pick up the geometrical interpretation for 1,2, and 3 space dimensions, you can move a little faster so you can make up some time. My first exam covers up through 10.4 of Stewart at the end of the fourth week. Usually, by then I am well into lines and planes and beyond.

A SERIES OF CLASS NOTES TO INTRODUCE LINEAR AND NONLINEAR PROBLEMS TO ENGINEERS, SCIENTISTS, AND APPLIED MATHEMATICIANS

LINEAR CLASS NOTES:
A COLLECTION OF HANDOUTS FOR
REVIEW AND PREVIEW
OF LINEAR THEORY
INCLUDING FUNDAMENTALS OF
LINEAR ALGEBRA

CHAPTER 1

Review of Elementary Algebra and Geometry in One, Two, and Three Dimensions

As engineers, scientists, and applied mathematicians, we are interested in geometry for two reasons:

1. The engineering and science problems that we are interested in solving involve time and space. We model time by $\mathbf{R}$ and (three dimensional) space using $\mathbf{R}^3$. However some problems can be idealized as one or two dimensional problems. i.e., using $\mathbf{R}$ or $\mathbf{R}^2$. We assume familiarity with $\mathbf{R}$ as a model for one dimensional space (Descartes) and with $\mathbf{R}^2$ as a model for two dimensional space (i.e., a plane). We assume less familiarity with $\mathbf{R}^3$ as a model of three space, but we do assume some.
2. Even if the state variables of interest are not position in space (e.g., amounts of chemicals in a reactor, voltages, and currents), we may wish to draw graphs in two or three dimensions that represent these quantities. Visualization, even of things that are not geometrical, may provide some understanding of the process.

It is this second reason (i.e., that the number of state variables in an engineering system of interest may be greater than three) that motivates us to consider "**vector spaces**" of dimension greater than three. When speaking of a system, the term "**degrees of freedom**" is also used, but we will usually use the more traditional term "**dimension**" in a problem solving setting where no particular application has been selected.

This leads to several definitions for the word "vector";

1. A physical quantity having magnitude and direction. (What your physics teacher told you.)
2. A directed line segment in the plane or in 3-space.
3. An element in $\mathbf{R}^2$ or $\mathbf{R}^3$ (i.e., an ordered pair or an ordered triple of real numbers).
4. An element in $\mathbf{R}^n$ (i.e., an ordered n-tuple of real numbers).
5. An element in $\mathbf{K}^n$ where $\mathbf{K}$ is a field (an abstract algebraic structure e.g., $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{C}$).
6. An element of a vector space (another abstract algebraic structure).

The above considerations lead to two separate topics concerning "vectors".

1. Linear Algebra.
2. Vector Analysis

In **Linear Algebra** a "vector" in an n dimensional **vector space** may represent n state variables and thus a system having n **degrees of freedom**. Although pictures of such vectors in one, two, and three dimensions may provide insight into the behavior of our system, these graphs need have no geometrical meaning. A system having n degrees of freedom and requiring n state variables resides in an n dimensional vector space which has no real geometrical meaning (e.g., amounts of n different chemicals in a reactor, n voltages or currents, tension in n different elements in a truss, flow rates in n different connected pipes, or amounts of money in n different accounts) . This does not diminish the usefulness of these vector spaces in solving engineering, science, and economic problems. A typical course in linear algebra covers topics in three areas:

1. Matrix Algebra.
2. Abstract Linear Algebra (Vector Space Theory).
3. How to Solve a System of m Linear Algebraic Equations in n Unknowns.

Although a model for the equilibrium or steady state of a linear system with n state variables usually has n equations in n unknowns, little is gained and much is lost by a restriction to n equations in n unknowns. Matrix algebra and abstract linear algebra (or vector space theory) provide the tools for understanding the methods used to solve m equations in n unknowns. The big point is that **no geometrical interpretation of a vector is needed to make vectors in linear algebra useful for solving engineering, scientific, and economic problems**. A geometrical interpretation is useful in one, two and three dimensions to help you understand what can happen and the "why's" of the solution procedure, but **no geometrical interpretation of a vector in "n" space is needed to make vectors in linear algebra useful for solving engineering, scientific, and economic problems**.

In **Vector Analysis** (Vector Calculus and Tensor Analysis) we are very much interested in the geometrical interpretation of our vectors (and tensors). Providing a mathematical model of geometrical concepts is the central theme of the subject. Interestingly, the "vector" spaces of interest from a linear algebra perspective are usually **infinite dimensional** as we are interested in physical quantities at **every point in physical space** and hence have an infinite number of degrees of freedom and hence an infinite number of state variables for the system. These are often represented using functions rather than n-tuples of numbers as well as "vector" valued functions (e.g., three ordered functions of time and space) . We divide our system models into two categories based on the number of variables needed to describe a state for our system:

1. **Discrete System Models** having only a finite number of state variables. These are sometimes referred to as **lumped parameter systems** in engineering books. Examples are electric circuits, systems of springs and "stirred" chemical reactors.
2. **Continuum System Models** having an infinite number of state variables in one , two, or three physical dimensions. These are sometimes referred to as **distributed parameter systems** in engineering books. Examples are Electromagnetic Theory (Maxwell's equations), Elasticity, and Fluid Flow (Navier-Stokes equations).

One last point. Although the static or equilibrium problem for a discrete or lumped parameter problem requires you to find values for n unknown variables, the dynamics problem requires values for these n variables for all time and hence requires an infinite number of values. Again, these are usually given by n functions of time or equivalently by an n-dimensional time varying "vector". (Note that here the term vector means n-tuple. When we use "vector" to mean n-tuple, rather than an element in a vector space, we will put it in quotation marks.)

Handout #2 FUNDAMENTAL PROPERTIES OF THE REAL NUMBER SYSTEM Professor Moseley

Geometrically, the real numbers **R** can be put in one-to-one correspondence with a line in space (Descartes). We call this the **real number line** and view **R** as representing one dimensional space. We also view **R** as being a fundamental model for time. However, we wish to also view **R** mathematically as a set with algebraic and analytic properties. In this approach, we postulate the existence of the set **R** and use axioms to establish the fundamental properties. Other properties are then developed using a theorem/proof/definition format. This removes the need to view **R** geometrically or temporally and allows us to algebraically move beyond three dimensions. **Degrees of freedom** is perhaps a better term, but **dimension** is standard in mathematics and we will use it. We note that once a mathematical mode has been developed, although sometimes useful for intuitive understanding, **no geometrical interpretation of R is needed in the solution process for engineering, scientific, and economic problems**. We organize the fundamental properties of the real number system into several groups. The first three are the standard **axiomatic properties** of **R**:

1) The **algebraic** (**field**) properties. (An **algebraic field** is an abstract algebraic structure. Examples are **Q**, **R**, and **C**. However, **N** and **Z** are not)
2) The **order** properties. (These give **R** its one dimensional nature.)
3) The **least upper bound property**. (This leads to the **completeness property** that insures that **R** has no "holes" and may be the hardest property of **R** to understand. **R** and **C** are complete, but **Q** is not complete)

All other properties of the real numbers follow from these axiomatic properties. Being able to "see" the geometry of the real number line may help you to intuit other properties of **R**. However, this ability is not necessary to follow the axiomatic development of the properties of **R**, and can sometimes obscure the need for a proof of an "obvious" property.

We consider other properties that can be derived from the axiomatic ones. To solve **equations** we consider

4)Additional algebraic (field) properties of **R** including **factoring** of polynomials.

To solve **inequalities**, we consider:

5) Additional order properties including the definition and properties of the **order relations** $<$, $>$, $\leq$, and $\geq$.

A **vector space** is another abstract algebraic structure. Application problems are often formulated and solved using the algebraic properties of fields and vector spaces. **R** is not only a field but also a (one dimensional) vector space. The last two groups of properties of interest show that **R** is not only a field and a vector space but also an **inner product space** and hence a **normed linear space**:

6) The **inner product** properties and
7) The **norm** (**absolute value or length**) properties.

These properties are of interest since the notion of length (norm) implies the notion of **distance apart** (**metric**), and hence provides topological properties and allows for **approximate solutions**. They can be algebraically extended to two dimensions ($\mathbf{R}^2$), three dimensions ($\mathbf{R}^3$), n dimensions ($\mathbf{R}^n$), and even infinite dimensions (**Hilbert Space**). There are many applications for vector spaces of dimension greater than three (e.g., solving m linear algebraic equations in n unknown variables) and infinite dimensional spaces (e.g., solving differential equations). In these applications, the term dimension does not refer to space or time, but to the **degrees of freedom** that the problem has. A physical system may require a finite or infinite number of **state variables** to specify the state of the system. Hence it would aid our intuition to replace the term **dimension** with the term **degrees of freedom**. On the other hand (OTOH), just as the real number line aids (and sometimes confuses) our understanding of the real number system, vector spaces in one two, and three dimensions can aid (and confuse) our understanding of the **algebraic** properties of vector spaces.

ALGEBRAIC (FIELD) PROPERTIES. Since the set **R** of real numbers along with the **binary operations** of addition and multiplication satisfy the following properties, the system denoted by the 5-tuple ($\mathbf{R}$, +, $\cdot$.0,1) is an example of a **field** which is an abstract algebraic structure.

A1.	$x+y = y+x \quad \forall x,y \in \mathbf{R}$	Addition is Commutative
A2.	$(x+y)+z = x+(y+z) \quad \forall x,y\ z \in \mathbf{R}$	Addition is Associative
A3.	$\exists\ 0 \in \mathbf{R}$ s.t. $x+0 = x \quad \forall x \in \mathbf{R}$	Existence of a Right Additive Identity
A4.	$\forall x \in \mathbf{R}\ \exists\ w$ s.t. $x+w = 0$	Existence of a Right Additive Inverse
A5.	$xy = yx \quad \forall x,y \in \mathbf{R}$	Multiplication is Commutative
A6.	$(xy)z = x(yz) \quad \forall x,y\ z \in \mathbf{R}$	Multiplication is Association
A7.	$\exists\ 1 \in \mathbf{R}$ s.t. $x \cdot 1 = x \quad \forall x \in \mathbf{R}$	Existence of a Right Multiplicative Identity
A8.	$\forall x \in \mathbf{R}$ s.t. $x\neq 0\ \exists\ w \in \mathbf{R}$ s.t. $xw=1$	Existence of a Right Multiplicative Inverse
A9.	$x(y+z) = xy + xz \quad \forall x,y\ z \in \mathbf{R}$	Multiplication Distributes over Addition

There are other algebraic (field) properties which follow from these nine fundamental properties. Some of these additional properties (e.g., cancellation laws) are listed in high school algebra books and calculus texts. Check out the above properties with specific values of x, y, and z. For example, check out property A3 with $x = \pi$. Is $\pi + 0 = \pi$?

ORDER PROPERTIES. There exists the subset $P_{\mathbf{R}} = \mathbf{R}^{+}$ of positive real numbers that satisfies the following:

O1. $x,y \in P_{\mathbf{R}}$ implies $x+y \in P_{\mathbf{R}}$

O2. $x,y \in P_{\mathbf{R}}$ implies $xy \in P_{\mathbf{R}}$

O3. $x \in P_{\mathbf{R}}$ implies $-x \notin P_{\mathbf{R}}$

O4. $x \in \mathbf{R}$ implies exactly one of $x = 0$ or $x \in P_{\mathbf{R}}$ or $-x \in P_{\mathbf{R}}$ holds (trichotomy).

Note that the order properties involve the binary operations of addition and multiplication and are therefore linked to the field properties. There are other order properties which follow from the four fundamental properties. Some of these are listed in high school algebra books and calculus texts. The symbols <, >, $\leq$, and $\geq$ can be defined using the set $P_{\mathbf{R}}$. For example, $a\leq b$ if (and only if) $b-a\in P_{\mathbf{R}}\cup\{0\}$. The properties of these relations can then be established. Check out the above properties with specific values of x and y. For example, check out property O1 with $x = \sqrt{2}$ and $y = \pi$. Is $\sqrt{2} + \pi$ positive?

LEAST UPPER BOUND PROPERTY: The **least upper bound property** leads to the **completeness property** that assures us that the real number line has no holes. This is perhaps the most difficult concept to understand. It means that we must include the irrational numbers (e.g. π and $\sqrt{2}$) with the rational numbers (fractions) to obtain all of the real numbers.

DEFINITION. If S is a set of real numbers, we say that b is an **upper bound** for S if for each $x \in S$ we have $x \leq b$. A number c is called a **least upper bound** for S if it is an upper bound for S and if $c \leq b$ for each upper bound b of S.

LEAST UPPER BOUND AXIOM. Every nonempty set S of real numbers that has an upper bound has a least upper bound.

Check this property out with the set $S = \{x \in \mathbf{R}: \ x^2 < 2\}$. Give three different upper bounds for the set S. What real number is the least upper bound for the set S? Is it in the set S?

ADDITIONAL ALGEBRAIC PROPERTIES AND FACTORING. Additional algebraic properties are useful in solving for the **roots** of the **equation** $f(x) = 0$ (i.e., the **zeros** of f) where f is a real valued function of a real variable. Operations on equations can be defined which result in **equivalent equations** (i.e., ones with the same solution set). These are called **equivalent equation operations** (EEO's). An important property of **R** states that if the product of two real numbers is zero, then at least one of them is zero. Thus if f is a polynomial that can be factored so that $f(x) = g(x)\,h(x) = 0$, then either $g(x) = 0$ or $h(x) = 0$ (or both since in logic we use the **inclusive or**). Since the degrees of g and h will both be less than that of f, this reduces a hard problem to two easier ones. When we can repeat the process to obtain a product of linear factors, we can obtain all of the zeros of f (i.e., the roots of the equation).

ADDITIONAL ORDER PROPERTIES. Often application problems require the solution of inequalities. The symbols $<$, $>$, $\leq$, and $\geq$ can be defined in terms of the set of positive numbers $P_{\mathbf{R}} = \mathbf{R}^+ = \{x \in \mathbf{R}: x > 0\}$ and the binary operation of addition. As with equalities, operations on inequalities can be developed that result in **equivalent inequalities** (i.e., ones that have the same solution set). These are called **equivalent inequality operations** (EIO's).

ADDITIONAL PROPERTIES. There are many additional properties of **R**. We consider here only the important absolute value property which is basically a function.

$$|x| = \begin{cases} -x & \text{if } x < 0 \\ x & \text{if } x \geq 0 \end{cases}$$

This function provides a norm, a metric, and a topology for **R**. Thus we can define limits, derivatives and integrals over intervals in **R**.

Handout #3 INTRODUCTION TO ALGEBRAIC STRUCTURES: Prof. Moseley
AN ALGEBRAIC FIELD

To introduce the notion of an **abstract algebraic structure** we consider (algebraic) **fields**. (These should not to be confused with vector and scalar fields in vector analysis.) Loosely, an algebraic field is a number system where you can add, subtract, multiply, and divide (except for dividing by zero of course). Examples of fields are **Q**, **R**, and **C**. Examples of other abstract algebraic structures are: Groups, Rings, Integral Domains, Modules and Vector Spaces.

DEFINITION. Let F be a set (of numbers) together with two **binary operations** (which we call addition and multiplication), denoted by + and $\cdot$ (or juxtaposition), which satisfy the following list of properties.:

F1) $\forall$ x,y $\in$ F $x + y = y + x$ Addition is commutative
F2) $\forall$ x,y $\in$ F $x + (y + z) = (x + y) + z$ Addition is associative
F3) $\exists$ an element $0 \in F$ such that. $\forall$ x $\in$ F, $x + 0 = x$ Existence of a right additive identity
F4) $\forall$ x $\in$ F, $\exists$ a unique $y \in F$ s.t. $x + y = 0$ Existence of a right additive inverse for each element
We usually denote y by -x for each x in F.
F5) $xy = yx$ $\forall$ x,y $\in$ F Multiplication is commutative
F6) $x(yz) = (xy)z$ $\forall$ x,y,z $\in$ F Multiplication is associative
F7) $\exists$ an element $1 \in F$ such that $1\neq0$ and $\forall$ x $\in$ F, $x1 = x$ Existence of a right multiplicative identity
F8) $\forall$ x s.t. $x\neq0$, $\exists$ a unique $y \in F$ s.t. $xy = 1$ Existence of a right multiplicative inverse for each element except 0
We usually denote y by (x^{-1}) or $(1/x)$ for each nonzero x in F.
F9) $x(y + z) = xy + xz$ $\forall$ x,y,z $\in$ F (Multiplication distributes over addition)

Then the ordered 5-tuple consisting of the set F and the structure defined by the two operations of addition and multiplication as well as two identity elements mentioned in the definition, $\mathsf{K} = (F,+,\cdot,0,1)$, is an algebraic **field**.

Although technically not correct, we often refer to the set F as the (algebraic) field. The elements of a field (i.e. the elements in the set F) are often called **scalars**. Since the letter F is used a lot in mathematics, the letter K is also used for a field of scalars. Since the rational numbers **Q**, the real numbers **R**, and the complex numbers **C** are examples of fields, it will be convenient to use the notation **K** to represent (the set of elements for) any field $\mathsf{K} = (\mathbf{K},+,\cdot,0,1)$.

The properties in the definition of a field constitute the fundamental axioms for **field theory**. Other field properties can be proved based only on these properties. Once we proved that **Q**, **R**, and **C** are fields (or believe that someone else has proved that they are) then, by the **principle of abstraction**, we need not prove these properties for these fields individually, one proof has done the work of three. In fact, for every (concrete) structure that we can establish as a field, all of the field properties apply.

We prove an easy property of fields. As a **field property**, this property must hold in every algebraic field. It is an identity and we use the standard form for proving identities.

THEOREM #1. Let $\mathsf{K} = (\mathbf{K},+,\cdot,0,1)$ be a field.. Then $\forall x\in\mathbf{K}$, $0 + x = x$. (In English, this property states that the right identity element 0 established in F3 is also a left additive identity element.)
Proof. Let $x\in\mathbf{K}$. Then

STATEMENT	REASON
$0 + x = x + 0$	F1. Addition is commutative
$= x$	F3 0 is a rightr additive identity element. (and properties of equality)

Hence $\forall x \in \mathbf{K}$, $0 + x = x$.

Q.E.D.

Although this property appears "obvious" (and it is) a formal proof can be written. This is why mathematicians may disagree about what axioms to select, but once the selection is made, they rarely disagree about what follows logically from these axioms. Actually, the first four properties establish a field with addition as an **Abelian (or commutative) group** (another abstract algebraic structure). **Group Theory**, particularly for Abelian groups have been well studied by mathematicians. Also, if we delete 0, then the nonzero elements of **K** along with multiplication also form an Abelian group. We list some (Abelian) group theory properties as they apply to fields. Some are identities, some are not.

THEOREM #2. Let $\mathcal{K} = (\mathbf{K},+,\cdot,0,1)$ be a field. Then
1. The identity elements 0 and 1 are unique.
2. For each nonzero element in **K**, its additive and multiplicative inverse element is unique.
3. 0 is its own additive inverse element (i.e., $0 + 0 = 0$) and it is unique, but it has no multiplicative inverse element.
4. The additive inverse of an additive inverse element is the element itself. (i.e., if $-a$ is the additive inverse of a, then $-(-a) = a$).
5. $-(a + b) = -a + -b$. (i.e., the additive inverse of a sum is the sum of their additive inverses.)
6. The multiplicative inverse of a multiplicative inverse element is the element itself. (i.e., if a^{-1} is the multiplicative inverse of a, then $(a^{-1})^{-1} = a$).
7. $(a\,b)^{-1} = a^{-1}\,b^{-1}$. (i.e., the multiplicative inverse of a product is the product of their multiplicative inverses.)
8. Sums and products can be written in any order you wish.
9. If $a + b = a + c$, then $b = c$. (Left Cancellation Law for Addition)
10. If $ab = ac$ and $a \neq 0$, then $b = c$. (Left Cancellation Law for Multiplication.)

Traditionally the concept of a **vector** is introduced by physicists as a quantity such as force or velocity that has magnitude and direction. Vectors are then represented geometrically by **directed line segments** and are thought of geometrically. Modern mathematical treatments introduce vectors algebraically as elements in a **vector space**. As an introductory compromise we introduce two dimensional vectors algebraically and then examine the correspondence between algebraic vectors and directed line segments in the plane. To define a **vector space** algebraically, we need a set of vectors and a set of scalars. We also need to define the algebraic operations of **vectors addition** and **scalar multiplication**. Since we associate the analytic set $\mathbf{R}^2 = \{(x,y): x,y \in \mathbf{R}\}$ with the geometric set of points in a plane, we will use another notation, $\mathbb{R}^2$, for the set of two dimensional **algebraic** vectors.

SCALARS AND VECTORS. We define our set of scalars to be the set of real numbers $\mathbf{R}$ and our set of vectors to be the set of ordered pairs $\mathbb{R}^2 = \{[x,y]^T: x,y \in \mathbf{R}\}$. We use the matrix notation $[x,y]^T$ (i.e.[x,y] transpose, see Chapter 2) to indicate a **column vector** to save space. (We could use row vectors, [x,y], but when we solve liear algebraqic equations we prefer column vectors.) When writing homework papers it is better to use column vectors explicitly. We write $\vec{x} = [x,y]^T$ and refer to x and y as the **components** of the vector $\vec{x}$.

VECTOR ADDITION. We define the **sum** of the vectors $\vec{x}_1 = [x_1,y_1]^T$ and $\vec{x}_2 = [x_2,y_2]^T$ as the vector $\vec{x}_1 + \vec{x}_2 = [x_1 + x_2, y_1 + y_2]^T$. For example, $[2,1]^T + [5,-2]^T = [7,-1]^T$. Thus we add vectors component wise.

SCALAR MULTIPLICATION. We define **scalar multiplication** of a vector $\vec{x} = [x,y]^T$ by a scalar $\alpha \in \mathbf{R}$ by $\alpha\,\vec{x} = [\alpha\, x, \alpha\, y]^T$. For example, $3[2,-1]^T = [6,-3]^T$. Thus we multiply each component in $\vec{x}$ by the scalar α. We define $\hat{i} = [1, 0]^T$ and $\hat{j} = [0, 1]^T$ so that every vector $\vec{x} = [x, y]^T$ may be written uniquely as $\vec{x} = x\,\hat{i} + y\,\hat{j}$.

GEOMETRICAL INTERPRETATION. Recall that we associate the analytic set $\mathbf{R}^2 = \{(x,y): x,y \in \mathbf{R}\}$ with the geometric set of points in a plane. Temporarily, we use $\tilde{x} = (x,y)$ to denote a point in $\mathbf{R}^2$. We might say that the points in a plane are a geometric interpretation of $\mathbf{R}^2$. We can establish a **one-to-one correspondence** between the analytic set $\mathbf{R}^2$ and geometrical vectors (**directed line segments**). First consider only **directed line segments** which are **position vectors**; that is, have their "tails" at the origin (i.e. at the point $\tilde{0} = (0,0)$ and their heads at some other point, say, the point $\tilde{x} = (x,y)$ in the plane $\mathbf{R}^2 = \{(x,y): x, y \in \mathbf{R}\}$. Denote this set by $\mathbf{G}$.

Position vectors are said to be "based at $\vec{0}$". If $\tilde{x} \in \mathbf{R}^2$ is a point in the plane, then we let $\overrightarrow{0\,x}$ denote the position vector from the origin $\tilde{0}$ to the point $\tilde{x}$. "Clearly" there exist a one-to-one correspondence between $\mathbf{G}$ and the set $\mathbf{R}^2$ of points in the plane; that is, we can readily identify exactly one vector in $\mathbf{G}$ with exactly one point in $\mathbf{R}^2$. Now "clearly" a one-to-one correspondence also exists between the set of points in $\mathbf{R}^2$ and the set of algebraic vectors in $\mathbb{R}^2$. Hence a one-to-one correspondence exists between the set of algebraic vectors $\mathbb{R}^2$ and the set of geometric vectors $\mathbf{G}$. In a traditional treatment, vector addition and scalar multiplication are defined geometrically for directed line segments in $\mathbf{G}$. We must then prove that the geometric definition of the addition of two directed line segments using the **parallelogram law** corresponds to algebraic addition of the corresponding vectors in $\mathbb{R}^2$. Similarly for scalar multiplication. We say that we must prove that the two structures are **isomorphic**. It is somewhat simpler to define vector addition and scalar multiplication algebraically on $\mathbb{R}^2$ and think of **G** as a geometric interpretation of the two dimensional vectors. One can then develop the **theory of vectors** without getting bogged down in the geometric proofs to establish the isomorphism. We can extend this isomorphism to $\mathbf{R}^2$. However, although we readily accept adding geometric vectors in **G** and will accept with coaxing adding algebraic vectors in $\mathbb{R}^2$, we may baulk at the idea of adding points in $\mathbf{R}^2$. However, the distinction between these sets with structure really just amounts to an interpretation of ordered pairs rather than an inherent difference in the mathematical objects being considered. Hence from now on, for any of these we will use the symbol $\mathbf{R}^2$ unless there is a philosophical reason to make the distinction. Reviewing, we have

$$\mathbb{R}^2 \cong \mathbf{G} \cong \mathbf{R}^2$$

where we have used the symbol $\cong$ to denote that there is an isomorphism between these sets with structure.

MAGNITUDE AND DIRECTION. Normally we think of directed line segments as being "free" vectors; that is, we identify any directed line segment in the plane with the directed line segment in **G** which has the same **magnitude** (length) and **direction**. The magnitude of the algebraic vector $\vec{x} = [x,y]^T \in \mathbb{R}^2$ is defined to be $||\vec{x}|| = \sqrt{x^2 + y^2}$ which, using Pythagoras, is the length of the directed line segment in **G** which is associated with $\vec{x}$. It is easy to show that the property (i.e. prove the theorem) $||\alpha\,\vec{x}|| = |\alpha|\ ||\vec{x}||$ holds for all $\vec{x} \in \mathbb{R}^2$ and all scalars α. It is also easy to show that if $\vec{x} \neq 0$ the vector $\vec{x}\,/||\vec{x}|| = (1/||\vec{x}||)\ \vec{x}$ has magnitude equal to one (i.e. is a **unit vector**). Examples of unit vectors are $\hat{i} = [1, 0]^T$ and $\hat{j} = [0, 1]^T$. "Obviously" any nonzero vector $\vec{x} = [x,y]^T$ can be written as $\vec{x} = ||\vec{x}||\ \hat{u}$ where $\hat{u} = \vec{x}\,/||\vec{x}||$ is a unit vector in the same direction as $\vec{x}$. That is, $\hat{u}$ gives the direction of $\vec{x}$ and $||\vec{x}||$ gives its magnitude. For example, $\vec{x} = [3, 4]^T = 5\,[3/5, 4/5]^T$ where $\hat{u} = [3/5, 4/5]^T$ and $||\vec{x}|| = 5$. To make vectors in $\mathbb{R}^2$ more applicable to two dimensional geometry we can introduce the concept of an **equivalence relation** and **equivalence classes**. We say that an arbitrary directed line segment in the plane is **equivalent** to a geometrical vector in **G** if it has the same direction and magnitude. The set of all directed line segments equivalent to a given vector in **G** forms an equivalence class. Two directed line segments are **related** if they are in the same equivalence classes. This relation is called an **equivalence relation** since all directed line segments that are related can be thought of as being the same. The equivalence classes partition the set of all directed line segments into sets that are mutually exclusive whose union is all of the directed line segments.

VECTORS IN $\mathbf{R}^3$. Having established an isomorphism between $\mathbf{R}^2$, $\mathbb{R}^2$, and G, we make no distinction in the future and will usually use $\mathbf{R}^2$ for the set of vectors in the plane, the set of points in the plane and the set of geometrical vectors. The context will explain what is meant. A similar development can be done algebraically and geometrically for vectors in 3-space. There is a technical difference between the sets $\mathbf{R}\times\mathbf{R}\times\mathbf{R} = \{(x,y,z): x, y, z\in\mathbf{R}\}$, $\mathbf{R}^2\times\mathbf{R} = \{((x,y),z): x, y, z\in\mathbf{R}\}$, and $\mathbf{R}\times\mathbf{R}^2 = \{(x,(y,z)): x, y, z\in\mathbf{R}\}$, but they are all isomorphic and we will usually consider them to be the same and denote all three by $\mathbf{R}^3$. Furthermore, similar to 2 dimensions, we will use $\mathbf{R}^3$ to represent: the set of **ordered triples**, the set of points in 3 dimensional space, the set of 3 dimensional position vectors given by $\vec{x}$ = $[x,y,z]^T = x\hat{i} + y\hat{j} + z\hat{k}$ where $\hat{i} = [1, 0, 0]^T$, $\hat{j}$ = $[0, 1, 0]^T$, and $\hat{k}$ = $[0, 0, 1]^T$, and the set of any three dimensional geometric vectors denoted by directed line segments.

The same analytic (but not geometric) development can be done for 4,5,..., n to obtain the set $\mathbf{R}^n$ of n-dimensional vectors. Again, no "real" geometrical interpretation is available. This does not diminish the usefulness of **n-dimensional space** for engineering since, for example, it plays a central role in the theory behind solving linear algebraic equations (m equations in n unknowns) and the theory behind any system which has n state variables (e.g., springs, circuits, and stirred tank reactors). The important thing to remember is that, although we may use geometric language for $\mathbf{R}^n$, we are doing algebra (and/or analysis), and not geometry.

PROPERTIES THAT CAN BE EXTENDED TO HIGHER DIMENSIONS. There are additional algebraic, analytic, and topological properties for $\mathbf{R}$, $\mathbf{R}^2$, and $\mathbf{R}^3$ that deserve to be listed since they are easily extended from $\mathbf{R}$, $\mathbf{R}^2$,$\mathbf{R}^3$ to $\mathbf{R}^4$,..., $\mathbf{R}^n$,... and, indeed to the (countably) infinite dimensional Hilbert space. We list these in four separate categories: inner product, norm, metric, and topology. Topologists may start with the definition of a topology as a collection of open sets. However, we are only interested in topologies that come from metrics which come from norms.

INNER (DOT) PRODUCT PROPERTIES. For $x_1, x_2 \in \mathbf{R}$, define $<x_1,x_2> = x_1x_2$. For $\vec{x}_1$ = $[x_1,y_1]^T = x_1\hat{i} + y_1\hat{j}$, $\vec{x}_2 = [x_2,y_2]^T = x_2\hat{i} + y_2\hat{j} \in \mathbf{R}^2$, define $<\vec{x}_1,\vec{x}_2> = x_1x_2 + y_1y_2$. For for $\vec{x}_1$ = $[x_1,y_1,z_1]^T = x_1\hat{i} + y_1\hat{j} + z_1\hat{k}$, $\vec{x}_2 = [x_2,y_2,z_2]^T = x_2\hat{i} + y_2\hat{j} + z_2\hat{k} \in \mathbf{R}^3$ define $<\vec{x}_1,\vec{x}_2> = x_1x_2 + y_1y_2 + z_1z_2$. Then in $\mathbf{R}$, $\mathbf{R}^2$, and $\mathbf{R}^3$ and, indeed, in any (real) inner product space V, we have the following (abstract) properties of an inner product:

IP1) $<\vec{x},\vec{y}> = <\vec{y},\vec{x}> \quad \forall \vec{x},\vec{y} \in V$

IP2) $<\alpha\,\vec{x} + \beta\,\vec{y}, \vec{z}> = \alpha <\vec{x},\vec{z}> + \beta <\vec{y},\vec{z}> \quad \forall x,y,z \in V, \quad \forall \alpha,\beta \in \mathbf{R}$

IP3) $<\vec{x},\vec{x}> \; > 0 \;$ if $\vec{x} \neq \vec{0}$

$<\vec{x},\vec{x}> \; = 0 \;$ if $\vec{x} = \vec{0}$

Check these properties out using specific values of x, y, z, α, β. For example, check out property IP2 with $\alpha = 2$, $\beta = 4$, $x = 1$, $y = 3$ and $z = 5$. First calculate the left hand side (LHS), then the right hand side (RHS). Are they the same? Is the property true for these values? Now do the same when $\vec{x}_1 = [2,3]^T = 2\hat{i} + 3\hat{j}$, $\vec{x}_2 = [1,4]^T = \hat{i} + 4\hat{j}$ and $\alpha = 2$ and $\beta = 3$. Also try some three dimensional vectors.

The inner or dot product of two vectors in two and three dimensions is more often denoted by $\vec{x}\cdot\vec{y}$. We have used the "physics" notation $<\vec{x},\vec{y}>$ in 2 and 3 dimensions since (x,y) is used to denote the coordinates of a point in the plane $\mathbf{R}^2$ and (x,y,z) for a point in 3 space. Note that we do not use arrows over the "vectors" when we are in one dimension. In $\mathbf{R}^2$ and $\mathbf{R}^3$, the geometric definition of inner (dot) product is $<\vec{x}_1,\vec{x}_2> = \|\vec{x}_1\|\,\|\vec{x}_2\|\cos\theta$ where $\|\cdot\|$ is the norm (magnitude) (see below) of the vector and θ is the angle between the two vectors. In $\mathbf{R}^n$, the definition of the inner product is $(\vec{x},\vec{y}) = \sum_{i=1}^{n} x_i y_i$ where $\vec{x} = (x_1, x_2, ..., x_n) = \{x_i\}_{i=1}^{n}$ and

$\vec{y} = \{y_i\}_{i=1}^{n}$. Note we have moved to algebraic notation (no geometric interpretation is available).

NORM (ABSOLUTE VALUE OR LENGTH) PROPERTIES. For $x \in \mathbf{R}$, define $\|x\| = \sqrt{< x, x >} = \sqrt{x \cdot x} = \sqrt{x^2} = |x|$, i.e., the absolute value of the number. For $\vec{x} = [x,y]^T = x\,\hat{i} + y\,\hat{j} \in \mathbf{R}^2$, define $\|\vec{x}\| = \sqrt{x^2 + y^2}$. For for $\vec{x} = [x,y,z]^T = x\,\hat{i} + y\,\hat{j} + z\,\hat{k}$ define $\|\vec{x}\| = \sqrt{x^2 + y^2 + z^2}$. Then in $\mathbf{R}$, $\mathbf{R}^2$, and $\mathbf{R}^3$ and, indeed, in any normed linear space V, we have the following (abstract) properties of a norm:

N1) $\|\vec{x}\| > 0 \quad \text{if} \quad \vec{x} \neq \vec{0}$

$\|\vec{x}\| = 0 \quad \text{if} \quad \vec{x} = \vec{0}$ = 0 if x = 0

N2) $\|\alpha\,\vec{x}\| = |\alpha|\,\|\vec{x}\| \quad \forall \vec{x} \in V, \quad \forall \alpha \in \mathbf{R}$

N3) $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\| \quad \forall \vec{x}, \vec{y} \in V \qquad \forall\, x, y \in \mathbf{R}$ (Triangle Inequality)

Check these properties out using specific values of x and α. For example check out property 2 with α = -3 and x = 2. That is, calculate the left hand side (LHS), then the right hand side (RHS). Are they the same? Is the property true for these values? Now check out these properties for two and three dimensional vectors.

In $\mathbf{R}$, $\mathbf{R}^2$ and $\mathbf{R}^3$, the norm is the geometric magnitude or length of the vector. In $\mathbf{R}^n$, the norm of the vector $\vec{x} = \{x_i\}_{i=1}^{n}$ is defined to be $\|\vec{x}\| = \sqrt{\sum_{i=1}^{n} x_i^{\,2}}$.

METRIC. For $x_1, x_2 \in \mathbf{R}$, define $\rho(x_1, x_2) = |x_1 - x_2|$. For $\vec{x}_1 = [x_1,y_1]^T = x_1\,\hat{i} + y_1\,\hat{j}$, $\vec{x}_2 = [x_2,y_2]^T = x_2\,\hat{i} + y_2\,\hat{j} \in \mathbf{R}^2$, define $\rho(\vec{x}_1, \vec{x}_2) = \|\vec{x}_1 - \vec{x}_2\| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$. For $\vec{x}_1 = [x_1,y_1,z_1]^T = x_1\,\hat{i} + y_1\,\hat{j} + z_1\,\hat{k}$, $\vec{x}_2 = [x_2,y_2,z_2]^T = x_2\,\hat{i} + y_2\,\hat{j} + z_2\,\hat{k} \in \mathbf{R}^3$ define $\rho(\vec{x}_1, \vec{x}_2) = \|\vec{x}_1 - \vec{x}_2\| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$. Then in $\mathbf{R}$, $\mathbf{R}^2$, and $\mathbf{R}^3$, and indeed, in any metric space V, we have the following (abstract) properties of a metric:

M1 $\rho(\vec{x}, \vec{y}) > 0 \quad \text{if} \quad \vec{x} \neq \vec{y}$

$\rho(\vec{x}, \vec{y}) = 0 \quad \text{if} \quad \vec{x} = \vec{y}$

M2 $\rho(\vec{x}, \vec{y}) = \rho(\vec{y}, \vec{x}) \quad \forall \vec{x}, \vec{y} \in V$

M3 $\rho(\vec{x}, \vec{z}) \leq \rho(\vec{x}, \vec{y}) + \rho(\vec{y}, \vec{z}) \quad \forall \vec{x}, \vec{y}, \vec{z} \in V$

In $\mathbf{R}$, $\mathbf{R}^2$ and $\mathbf{R}^3$, the metric is the geometric distance between the tips of the position vectors. In $\mathbf{R}^n$, the metric of the vectors $\vec{x} = \{x_i\}_{i=1}^{n}$, $\vec{y} = \{y_i\}_{i=1}^{n} \in \mathbf{R}^n$ is defined to be $\rho(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$. A metric yields a definition of limit. Although in $\mathbf{R}^n$ with n > 3, this in not geometric, it does provides a measure for evaluating approximate solutions to engineering problems having n state variables. This also allows for the concept of a sequence of approximate solutions that converges to the exact solution in engineering problems.

TOPOLOGY. A topology is a collection of (open) sets having certain properties. Intuitively, an open set is one that does not contain any of its boundary points. For example, open intervals are open sets in the usual topology for **R**. Disks that do not contain any boundary points are open sets in $\mathbf{R}^2$. Balls which do not contain any boundary points are open sets in $\mathbf{R}^3$. Similarly for $\mathbf{R}^n$, but no geometry. Topologists characterize continuity in terms of open sets.

Handout # 5 COMPLEX NUMBERS, GEOMETRY, AND COMPLEX FUNCTIONS OF A COMPLEX VARIABLE Prof. Moseley

To solve $z^2 + 1 = 0$ we "invent" the number i with the defining property $i^2 =$) 1. (Electrical engineers use j instead if i.) We then "define" the set of complex numbers as **C** = {x+iy:x,y∈**R**}. Then z = x+iy is known as the **Euler form** of z and z = (x,y)∈**C** = {(x,y):x,y∈**R**} is the **Hamilton form** of z where we identify **C** with $\mathbf{R^2}$ (i.e., with points in the plane). If $z_1 = x_1+iy$, and $z_2 = x_2+iy_2$, then $z_1+z_2 =_{df} (x_1+x_2) + i(y_1+y_2)$ and $z_1z_2 =_{df} (x_1x_2-y_1y_2) + i(x_1y_2+x_2y_2)$. Using these definitions, the nine properties of addition and multiplication in the definition of an abstract algebraic field can be proved. Hence the system (**C**, + ,·,0,1) is an example of an abstract algebraic field. Computation of the product of two complex numbers is made easy using the Euler form and your knowledge of the algebra of **R**, the FOIL (<u>f</u>irst, <u>o</u>uter, <u>i</u>nner, <u>l</u>ast) method, and the defining property $i^2 = -1$. Hence we have $(x_1+iy_1)(x_2+iy_2) = x_1x_2+x_1iy_2 + iy_1y_2 + i^2y_1y_2 = x_1x_2 + i\,(x_1y_2 + y_1y_2) - y_1y_2 = (x_1x_2 - y_1y_2) + i\,(x_1y_2+x_2y_1)$. This makes evaluation of polynomial functions easy.

<u>EXAMPLE #1</u>. If $f(z) = (3 + 2i) + (2 + i)z + z^2$, then

$$\begin{aligned} f(1+i) &= (3 + 2i) + (2 + i)(1 + i) + (1 + i)^2 = (3 + 2i) + (2 + 3i + i^2) + (1 + 2i + i^2) \\ &= (3 + 2i) + (2 + 3i - 1) + (1 + 2i - 1) = (3 + 2i) + (1 + 3i) + (\ 2i\ ) = \ 4 + 7i \end{aligned}$$

Division and evaluation of rational functions is made easier by using the **complex conjugate**. We also define the magnitude of a complex number as the distance to the origin in the complex plane.

<u>DEFINITION #1.</u> If z = x+iy, then the **complex conjugate** of z is given by $\overline{z} = x - iy$. Also the **magnitude** of z is $|z| = \sqrt{x^2 + y^2}$.

<u>THEOREM #1.</u> If $z,z_1,z_2 \in$ **C**, then a) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$, b) $\overline{z_1z_2} = \overline{z_1}\ \overline{z_2}$, c) $|z|^2 = z\,\overline{z}$, d) $\overline{\overline{z}} = z$.

<u>REPRESENTATIONS</u>. Since each complex number can be associated with a point in the (complex) plane, in addition to the **rectangular representation** given above, complex numbers can be represented using **polar coordinates**.

$$\begin{aligned} z &= x + iy = r\cos\theta + i\sin\theta = r\,(\cos\theta + i\sin\theta) \\ &= r\,\underline{/\ \theta\ } \text{ (polar representation)} \end{aligned}$$

Note that r = |z|. You should be able to convert from Rectangular form to Polar form and vice versa. For example, $2\,\underline{/\pi/4} = \sqrt{2} + \sqrt{2}$ i and $1 + \sqrt{3}\,i = 2\,\underline{/\,\pi/3}$. Also, if $z_1 = 3 + i$ and $z_2 = 1 + 2i$, then
$$\frac{z_1}{z_2} = \frac{z_1}{z_2}\frac{\overline{z_2}}{\overline{z_2}} = \frac{z_1\overline{z_2}}{|z_2|^2} = \frac{(3+i)(1+2i)}{(1+2^2)} = \frac{(3+6i+i+2i^2)}{(1+4)} = \frac{(3-2+7i)}{5} = \frac{1}{5} + \frac{7}{5}i$$

<u>EXAMPLE # 2</u> If $f(z) = \dfrac{z+(3+i)}{z+(2+i)}$, evaluate f(4+i).

Solution. $f(4+i)) = \dfrac{(4+i)+(3+i)}{(4+i)+(2+i)} = \dfrac{(7+2i)}{(6+3i)} = \dfrac{(7+2i)}{(6+3i)}\,\dfrac{(6-3i)}{(6-3i)}$

$$= \frac{(42+6)+(12-21)i}{(36+9)} = \frac{48-9i}{45} = \frac{48}{45} - \frac{9}{45}i = \frac{16}{15} - \frac{1}{5}i$$

<u>THEOREM #2</u>. If $z_1 = r_1\underline{\angle\theta_1}$ and $z_2 = r_2\underline{\angle\theta_2}$, then a) $z_1 z_2 = r_1 r_2 \underline{\angle\theta_1 + \theta_2}$, b) If $z_2 \neq 0$, then $\dfrac{z_1}{z_2} = \dfrac{r_1}{r_2}\underline{\angle\theta_1 - \theta_2}$, c) $z_1^2 = r_1^2\underline{\angle 2\theta_1}$, d) $z_1^n = r_1^n\underline{\angle n\theta_1}$.

<u>EULER'S FORMULA</u>. By definition $e^{i\theta} =_{df} \cos\theta + i\sin\theta$. This gives another way to write complex numbers in polar form.

$z = 1 + i\sqrt{3} = 2(\cos\pi/3 + i\sin\pi/3) = 2\,\underline{\angle\pi/3} = 2e^{i\pi/3}$

$z = \sqrt{2} + i\sqrt{2} = 2(\cos\pi/4 + i\sin\pi/4) = 2\,\underline{\angle\pi/4} = e^{i\pi/4}$

More importantly, it can be shown that this definition allows the extension of exponential, logarithmic, and trigonometric functions to complex numbers and that the standard properties hold. This allows you to determine what these extensions should be and to evaluate these functions.

<u>EXAMPLE #6</u>. If $f(z) = (2+i)\,e^{(1+i)z}$, find $f(1+i)$.
<u>Solution</u>. First $(1+i)(1+i) = 1 + 2i + i^2 = 1 + 2i - 1 = 2i$. Hence

$f(1+i) = (2+i)\,e^{2i} = (2+i)(\cos 2 + i\sin 2) = 2\cos 2 + i(\cos 2 + 2\sin 2) + i^2\sin 2$

$= 2\cos 2 - \sin 2 + i(\cos 2 + 2\sin 2)$ (exact answer)

$\approx 2(-0.4161468) - (0.9092974) + i(-0.4161468 + 2(0.9092974))$

$= -1.7415911 + i(1.4024480)$ (approximate answer)

How do you know that $-1.7415911 + i(1.4024480)$ is a good approximation to $2\cos 2 - \sin 2 + i(\cos 2 + 2\sin 2)$? Can you give an expression for the distance between these two numbers?

In the complex plane, e^z is not one-to-one. Restricting the domain of e^z to the strip $\mathbf{R}\times(-\pi,\pi]$) $\{(0,0)\}$, in the complex plane, we can define Ln z as the (compositional) inverse function of e^z with this restricted domain. This is similar to restricting the domain of sin x to $[-\pi/2, \pi/2]$ to obtain $\sin^{-1} x$ (Arcsin x) as its (compositional) inverse function.

<u>EXAMPLE #7</u>. If $f(z) = \text{Ln}\,[(1+i)z]$, find $f(1+i)$.
<u>Solution</u>. First $(1+i)(1+i) = 1 + 2i + i^2 = 1 + 2i - 1 = 2i$. Now let $w = x + iy = f(1+i)$ where $x, y \in \mathbf{R}$. Then $w = x + iy = \text{Ln}\,2i$. Since Ln z is the (compositional) inverse of the function of e^z, we have that $2i = \exp(x+iy) = e^x e^{iy} = e^x(\cos y + i\sin y)$. Equating real and imaginary parts we obtain $e^x \cos y = 0$ and $e^x \sin y = 2$ so that $\cos y = 0$ and hence $y = \pm\pi/2$. Now $e^x \sin\pi/2 = 2$ yields $x = \ln 2$, but $e^x \sin(-\pi/2) = 2$ implies $e^x = -2$. Since this is impossible, we have the unique solution $f(1+i) = \ln 2 + i(\pi/2)$. We check by direct computation using Euler's formula: $\exp[\ln 2 + i(\pi/2)] = e^{\ln 2}\, e^{i(\pi/2)} = 2(\cos(\pi/2) + i\sin(\pi/2)) = 2i$.
This is made easier by noting that if $z = x + iy = re^{i\theta}$, then $\text{Ln}\,z = \text{Ln}(r\,e^{i\theta}) = \ln r + i\theta$.

Letting $e^{iz} = \cos z + i\sin z$ we have $e^{-iz} = \cos z + i\sin(-z) = \cos z - i\sin z$. Adding we obtain $\cos z = (e^{iz} + e^{-iz})/2$ which extends cosine to the complex plane. Similarly $\sin z = (e^{iz} - e^{-iz})/2i$.

EXAMPLE #8. If $f(z) = \cos[(1 + i)z]$, find $f(1 + i)$.
Solution. First recall $(1 + i)(1 + i) = 1 + 2i + i^2 = 1 + 2i - 1 = 2i$. Hence
$f(1 + i) = \cos(2i) = (e^{i\,2i} + e^{i(-2i)})/2 = (e^{-2} + e^{2})/2 = \cosh 2$ (exact answer)
≈ 3.762195691. (approximate answer)
How do you know that 3.762195691 is a good approximation to cosh 2 ? Can you give an expression for the distance between these two numbers (i.e., cosh 2 and 3.762195691)?

It is important to note that we represent **C** geometrically as points in the (complex) plane. Hence the domain of f:**C**→**C** requires two dimensions. Similarly, the codomain requires two dimensions. Hence a "picture" graph of a complex valued function of a complex variable requires four dimensions. We refer to this a s a geometric depiction of f of Type 1. Hence we can not visualize complex valued function of a complex variable the same way that we do real valued function of a real variable. Again, geometric depictions of complex functions of a complex variable of Type 1 (picture graphs) are not possible. However, other geometric depictions are possible. One way to visualize a complex valued function of a complex variable is what is sometimes called a “set theoretic” depiction which we label as Type 2 depictions. We visualize how regions in the complex plane (domain) get mapped into other regions in the complex plane (codomain). That is, we draw the domain as a set (the complex z-plane) and then draw the w = f(z) plane. Often, points in the z-plane are labeled and their images in the w-plane are labeled with the same or similar notation. Another geometric depiction (of Type 3) is to let f(z) = f(x+iy) = u(x,y) + iv(x,y) and sketch the real and imaginary parts of the function separately.
Instead of drawing three dimensional graphs we can draw level curves which we call a Type 4 depiction. If we put these level curves on the same z-plane, we refer to it as a Type 5 depiction.

A SERIES OF CLASS NOTES TO INTRODUCE LINEAR AND NONLINEAR PROBLEMS TO ENGINEERS, SCIENTISTS, AND APPLIED MATHEMATICIANS

LINEAR CLASS NOTES:
A COLLECTION OF HANDOUTS FOR
REVIEW AND PREVIEW
OF LINEAR THEORY
INCLUDING FUNDAMENTALS OF
LINEAR ALGEBRA

CHAPTER 2

Elementary Matrix Algebra

Although usually studied together in a linear algebra course, **matrix algebra** can be studied separately from **linear algebraic equations** and **abstract linear algebra (Vector Space Theory)**. They are studied together since matrices are useful in the representation and solution of **linear algebraic equations** and yield important examples of the abstract algebraic structure called a **vector space**. Matrix algebra begins with the **definition of a matrix**. It is assumed that you have some previous exposure to matrices as an array of **scalars**. Scalars are elements of another abstract algebraic structure called a **field**. However, unless otherwise stated, the scalar entries in matrices can be assumed to be real or complex numbers (Halmos 1958,p.1). The definition of a matrix is followed closely by the definition of **basic algebraic operations** (computational procedures) which involve the scalar entries in the matrices (e.g., real or complex numbers) as well as possibly additional scalars. These include **unary**, **binary** or other types of operations.

Some basic operations which you may have previously seen are:

1. **Transpose** of a matrix.
2. **Complex Conjugate** of a matrix.
3. **Adjoint** of a matrix. (Not the same as the **Classical Adjoint** of a matrix.)
4. **Addition** of two matrices of a given size.
5. **Componentwise multiplication** of two matrices of a given size (not what is usually called multiplication of matrices).
6. Multiplication of a matrix (or a vector) by a scalar (**scalar multiplication**).
7. **Dot** (or **scalar**) **product** of two column vectors (or row vectors) to obtain a scalar. Also called an **inner product**. (Column and row vectors are one dimensional matrices.)
8. **Multiplication** (perhaps more correctly called composition) of two matrices with the right sizes (dimensions).

A **unary operation** is an operation which maps a matrix of a given size to another matrix of the same size. Unary operations are functions or mappings, but algebraists and geometers often think in terms of operating on an object to transform it into another object rather than in terms of mapping one object to another (already existing) object. Here the objects being mapped are not numbers, they are matrices. That is, you are given one matrix and asked to compute another. For square matrices, the first three operations, **transpose**, **complex conjugate**, and **adjoint**, are unary operations. Even if the matrices are not square, the operations are still **functions**.

A **binary operation** is an operation which maps two matrices of a given size to a third matrix of the same size. That is, instead of being given one matrix and being asked to compute another, you are given two matrices and asked to compute a third matrix. Algebraist think of combining or transforming two elements into a new element. **Addition** and **componentwise multiplication** (not what is usually called matrix multiplication), and **multiplication of square matrices** are examples of binary operations. Since on paper we write one element, the symbol for the binary operation, and then the second element, the operation may, but does not have to, depend on the order in which the elements are written. (As a mapping, the **domain** of a binary operation is a **cross product** of sets with elements that are **ordered pairs**.) A binary operation that does not depend on order is said to be **commutative**. Addition and componentwise multiplication are commutative; multiplication of square matrices is not.

This may be the first you have become aware that there are binary operations in mathematics that are not commutative. However, it is not the first time you have encountered a non-commutative binary operation. Instead of thinking of subtraction as the inverse operation of addition (which is effected by adding the additive inverse element), we may think of subtraction as a binary operation which is not commutative ($a - b$ is not usually equal to $b - a$). However, there is another fundamental binary operation on $\mathscr{F}(\mathbf{R},\mathbf{R})$ that is not commutative. If $f(x) = x^2$ and $g(x) = x + 2$, it is true that $(f+g)(x) = f(x) + g(x) = g(x) + f(x) = (g+f)(x)\ \forall x\in\mathbf{R}$ so that $f+g = g+f$. Similarly, $(fg)(x) = f(x)\,g(x) = g(x)\,f(x) = (gf)(x)\forall x\in\mathbf{R}$ so that $fg = gf$. However, $(f\circ g)(x) = f(g(x)) = f(x + 2) = (x + 2)^2$, but $(g\circ f)(x) = g(f(x)) = g(x^2) = x^2 + 2$ so that $f\circ g \neq g\circ f$. Hence the **composition of functions** is not commutative.

When multiplying a matrix by a scalar (i.e., **scalar multiplication**), you are given a matrix and a scalar, and asked to compute another matrix. When computing a **dot** (or **scalar**) **product**, you are given two column (or row) vectors and asked to compute a scalar.

EXERCISES on Introduction and Basic Operations

EXERCISE #1. True or False.

_____ 1. Computing the transpose of a matrix is a binary operation.
_____ 2. Multiplying two square matrices is a binary operation.
_____ 3. Computing the dot product of two column vectors is a binary operation.

Halmos, P. R.1958, *Finite Dimensional Vector Spaces* (Second Edition) Van Nostrand Reinhold Company, New York.

We begin with an informal definition of a matrix as an array of scalars. Thus an **m×n matrix** is an array of elements from an algebraic field **K** (e.g.,real or complex numbers) which has m rows and n columns. We use the notation $\mathbf{K}^{m\times n}$ for the set of all matrices over the field **K**. Thus, $\mathbf{R}^{3x4}$ is the set of all 3x4 real matrices and $\mathbf{C}^{2x3}$ is the set of all 2x3 matrices of complex numbers. We represent an arbitrary matrix over **K** by using nine elements:

$$\underset{mxn}{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdot & \cdot & \cdot & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdot & \cdot & \cdot & a_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{m,1} & a_{m,1} & \cdot & \cdot & \cdot & a_{m,n} \end{bmatrix} = [\, a_{ij} \,] \in \mathbf{K}^{mxn} \tag{1}$$

Recall that an **algebraic field** is an abstract algebra structure. **R** and **C** as well as **Q** are examples of algebraic fields. Nothing essential is lost if you think of the elements in the matrices as real or complex numbers in which case we refer to them as **real or complex matrices**. Parentheses as well as square brackets may be used to set off the array. However, the same notation should be used within a given discussion (e.g., a homework problem). That is, you must be consistent in your notation. Also, since vertical lines are used to denote the **determinant** of a matrix (it is assumed that you have had some exposure to determinants) they are not acceptable to denote matrices. (Points will be taken off if your notation is not consistent or if vertical lines are used for a matrix.) Since $\mathbf{K}^{mxn}$ for the set of all matrices over the field **K**, $\mathbf{R}^{mxn}$ is the set of all real matrices and $\mathbf{C}^{mxn}$ is the set of all complex matrices. In a similar manner to the way we embed **R** in **C** and think of the reals as a subset of the complex numbers, we embed $\mathbf{R}^{mxn}$ in $\mathbf{C}^{mxn}$ and, unless told otherwise, assume $\mathbf{R}^{mxn} \subseteq \mathbf{C}^{mxn}$. In fact, we may think of $\mathbf{C}^{mxn}$ as
$\mathbf{R}^{mxn} + i\,\mathbf{R}^{mxn}$ so that every complex matrix has a real and imaginary part. Practically, we simply expand the set where we are allowed to choose the numbers $a_{i\,j}$, i = 1, ..., m and j = 1, ..., n.

The elements of the algebraic field, $a_{i\,j}$, i = 1, ..., m and j = 1, ..., n, (e.g., real or complex numbers) are called the **entries** in (or **components** of) A. To indicate that $a_{i\,j}$ is the element in A is the i^{th} row and j^{th} column we have also used the shorthand notation $A = [a_{ij}]$. We might also use $A = (a_{ij})$. The first subscript of $a_{i\,j}$ indicates the row and the second subscript indicates the column. This is particularly useful if there is a formula for $a_{i\,j}$. We say that the **size** or **dimension** of the matrix is m×n and use the notation $\underset{mxn}{A}$ to indicate that A is a matrix of size m×n. The entries in A will also be called **scalars**. For clarity, it is sometimes useful to separate the subscripts with a comma instead of a space. For example we might use $a_{12,2}$ and $a_{i+3,i+j}$ instead of $a_{12\;2}$ and $a_{i+3\;i+j}$.

EXAMPLES. $\underset{3x3}{A} = \begin{bmatrix} 1 & 3/2 & 4 \\ 7 & 2 & 6 \\ 1 & 8 & 0 \end{bmatrix}$ and $B = (1/3, \quad 2/5, \quad 3)$ are examples of rational matrices.

$\underset{3x3}{A} = \begin{bmatrix} \pi & 3 & 4 \\ 7 & 2 & 6 \\ 1 & 8 & 0 \end{bmatrix}$, and $\underset{3x1}{B} = \begin{bmatrix} 1 \\ 4 \\ e \end{bmatrix}$ are examples of real (non-rational) matrices.

$A = (1+i, \quad 2, \quad 3+2i)$, and $\underset{3x2}{B} = \begin{bmatrix} \pi + 2i & e+i \\ 3/4 + 3i & 1.5 \\ 2e & 7\pi \end{bmatrix}$ are examples of complex (non-real) matrices.

As was done above, commas (instead of just spaces) are often used to separate the entries in $1\times n$ matrices (often called row vectors). Although technically incorrect, this practice is acceptable since it provides clarity.

COLUMN AND ROW VECTORS. An $m\times 1$ matrix (e.g., $\begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}$) is often called a **column vector**. We often use $[1,4,3]^T$ to denote this column vector to save space. The T stands for transpose as explained later. Similarly, a $1\times n$ matrix (e.g., [1, 4, 3]) is often called a **row vector**. This stems from their use in representing physical vector quantities such as position, velocity, and force. Because of the obvious isomorphisms, we will often use $\mathbf{R}^n$ for the set of real column vector $\mathbf{R}^{n\times 1}$ as well as for the set of real row vectors $\mathbf{R}^{1\times n}$. Similarly conventions hold for complex column and row vectors and indeed for $\mathbf{K}^n$, $\mathbf{K}^{n\times 1}$, and $\mathbf{K}^{1\times n}$. That is, we use the term "vector" even when n is greater than three and when the field is not $\mathbf{R}$. Thus we may use $\mathbf{K}^n$ for a finite sequence of elements in $\mathbf{K}$ of length n, no matter how we choose to write it.

EQUALITY OF MATRICES. We say that two matrices are **equal** if they are the same; that is, if they are identical. This means that they must be of the same size and have the same entries in corresponding locations. (We do not use the bracket and parenthesis notation within the same discussion. All matrices in a particular discussion must use the same notation.)

A FORMAL DEFINITION OF A MATRIX. Since $\mathbf{R}$ and $\mathbf{C}$ are examples of an (abstract) algebraic structure called a **field** (not to be confused with a vector field considered in vector analysis), we formally define a **matrix** as follows: An $m\times n$ **matrix** A over the (algebraic) field $\mathbf{K}$ (usually $\mathbf{K}$ is $\mathbf{R}$ or $\mathbf{C}$) is a function from the set $D = \{1,...,. m\} \times \{1, ..., n\}$ to the set $\mathbf{K}$ (i.e., $A: D \to \mathbf{K}$). D is the **cross product** of the sets $\{1,...,. m\}$ and $\{1, ..., n\}$. Its elements are **ordered pairs** (i,j) where $i\in\{1,...,. m\}$ and $j\in\{1, ..., n\}$. The value of the function A at the point $(i.j) \in D$ is denoted by a_{ij} and is in the codomain of the function. The value (i.j) in the domain that corresponds to a_{ij} is indicated by the location of a_{ij} in the array. Thus the array gives the values of the function at each point in the domain D. Thus the set of m×n matrices over the field K is the set of functions $\mathscr{F}(D,\mathbf{K})$ which we normally denote by $\mathbf{K}^{m\times n}$. The set of real m×n matrices is $\mathbf{R}^{m\times n} = \mathscr{F}(D,\mathbf{R})$ and the set of complex m×n matrices is $\mathbf{C}^{m\times n} = \mathscr{F}(D,\mathbf{C})$.

EXERCISES on Definition of a Matrix

<u>EXERCISE #1</u>. Write the 2×3 matrices whose entries are:
(a) $a_{ij} = i+j$, (b) $a_{ij} = (3i + j^2)/2$, (c) $a_{ij} = [\pi i + j]$.

<u>EXERCISE #2</u>. Write the 3×1 matrices whose entries are given by the formulas in Exercise 1.

<u>EXERCISE # 3</u>. Choose x and y so that A = B if:

(a) $\underset{1\times 3}{\mathbf{A}} = [x,2,3]$ and $\underset{1\times 3}{\mathbf{B}} = [4,y,3]$,

(b) $\underset{2\times 2}{\mathbf{A}_2} = \begin{bmatrix} 2 & x \\ 1 & 3 \end{bmatrix}$, $\underset{2\times 2}{\mathbf{B}} = \begin{bmatrix} 2 & 3 \\ y & 3 \end{bmatrix}$,

(c) $\underset{1\times 2}{\mathbf{A}} = [x,2]$, $\mathbf{B} = [y,2]$,

(d) $\underset{1\times 2}{\mathbf{A}} = [x,2]$, $\mathbf{B} = [3,1]$.

Let **K** be an arbitrary algebraic field. (e.g. **K** = **Q**, **R**, or **C**). Nothing essential is lost if you think of **K** as **R** or **C**. Recall that we use $\mathbf{K}^{m \times n}$ to denote the set of all matrices with entries in **K** that are of size (or dimension) mxn (i.e. having m rows and n columns). We first define the **transpose function** that can be applied to any matrix $A \in \mathbf{K}^{m \times n}$. It maps $\mathbf{K}^{m \times n}$ to $\mathbf{K}^{n \times m}$. We then define the **(complex) conjugate function** from $\mathbf{C}^{m \times n}$ to $\mathbf{C}^{m \times n}$. Finally, we define the **adjoint function** (not the same as the classical adjoint) from $\mathbf{C}^{m \times n}$ to $\mathbf{C}^{n \times m}$. We denote an arbitrary matrix A in $\mathbf{K}^{m \times n}$ by using nine elements.

$$\underset{m\times n}{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdot & \cdot & \cdot & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdot & \cdot & \cdot & a_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{m,1} & a_{m,1} & \cdot & \cdot & \cdot & a_{m,n} \end{bmatrix} = [\, a_{ij} \,] \in \mathbf{K}^{m \times n} \qquad (1)$$

TRANSPOSE.

DEFINITION #1. If $\underset{m\times n}{A} = [\, a_{ij} \,] \in \mathbf{K}^{m \times n}$, then $\underset{n\times m}{A^T} = [\, \tilde{a}_{ij} \,] \in \mathbf{K}^{n \times m}$ is defined by $\tilde{a}_{ij} = a_{j\,i}$ for i=1,...,n, j=1,...,m. A^T is called the **transpose** of the matrix A. (The transpose of A is obtained by exchanging its rows and columns.) The **transpose function** maps (or transforms) a matrix in $\mathbf{K}^{m \times n}$ to (or into) a matrix in $\mathbf{K}^{n \times m}$. Thus for the transpose function T, we have T: $\mathbf{K}^{m \times n} \to \mathbf{K}^{n \times m}$.

Note that we have not just defined one function, but rather an infinite number of functions, one for each set $\mathbf{K}^{m \times n}$. We have named the function T, but we use A^T instead of T(A) for the element in the codomain $\mathbf{K}^{n \times m}$ to which A is mapped. Note that, by convention, the notation $\underset{m\times n}{A} = [\, a_{ij} \,]$ means that a_{ij} is the entry in the i^{th} row and j^{th} column of A: that is, unless otherwise specified, the **order of the subscripts** (rather than the choice of index variable) indicates the row and column of the entry. Hence $\underset{n\times m}{A^T} \neq [\, a_{ji} \,]$.

EXAMPLE. If $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$, then $A^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$. If $B = \begin{bmatrix} -1 & i \\ 3+i & 2i \\ 5 & 1+i \end{bmatrix}$, the $B^T = \begin{bmatrix} -1 & 3+i & 5 \\ i & 2i & 1+i \end{bmatrix}$. If A is given by (1) above, then $A^T = \begin{bmatrix} a_{1,1} & a_{2,1} & \cdot & \cdot & \cdot & a_{n,1} \\ a_{1,2} & a_{2,2} & \cdot & \cdot & \cdot & a_{n,2} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{1,m} & a_{2,m} & \cdot & \cdot & \cdot & a_{n,m} \end{bmatrix}$. Note that this makes the rows and columns hard to follow.

PROPERTIES. There are lots of properties, particularly if $\mathbf{K} = \mathbf{C}$ but they generally involve other operations not yet discussed. Hence we give only one at this time. We give its proof using the standard format for proving identities in the STATEMENT/REASON format. We represent matrices using nine components, the first , second, and last components in the first, second, and last rows. We refer to this as the **nine element method**. For many proofs involving matrix properties, the nine element method is sufficient to provide a convincing argument.

THEOREM #1. For all $A \in \mathbf{K}^{m\times n}$ = the set of all m×n complex valued matrices, $(A^T)^T = A$.

Proof: Let

$$\underset{m\mathrm{x}n}{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdot & \cdot & \cdot & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdot & \cdot & \cdot & a_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{m,1} & a_{m,1} & \cdot & \cdot & \cdot & a_{m,n} \end{bmatrix} = [\, a_{ij} \,] \in \mathbf{K}^{m\mathrm{x}n} \tag{1}$$

Then by the definition of the transpose of A we have

$$\underset{m\mathrm{x}n}{A}^{T} = \begin{bmatrix} a_{1,1} & a_{2,1} & \cdot & \cdot & \cdot & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdot & \cdot & \cdot & a_{m,2} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{1,n} & a_{2,n} & \cdot & \cdot & \cdot & a_{m,n} \end{bmatrix}_{n\mathrm{x}m} = [\tilde{a}_{i,j}] \in \mathbf{K}^{n\mathrm{x}m} \tag{2}$$

where $\tilde{a}_{ji,j} = a_{j,i}$. We now show $(A^T)^T = A$ using the standard procedure for proving identities in the STATEMENT/REASON format.

<u>STATEMENT</u>	<u>REASON</u>
$\left\{\underset{m\mathrm{x}n}{A}^{T}\right\}^T = \left\{\begin{bmatrix} a_{1,1} & a_{1,2} & \cdot & \cdot & \cdot & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdot & \cdot & \cdot & a_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{m,1} & a_{m,2} & \cdot & \cdot & \cdot & a_{m,n} \end{bmatrix}^T\right\}^T$	Definition of A (using the nine element representation) (Notation)
$= \begin{bmatrix} a_{1,1} & a_{2,1} & \cdot & \cdot & \cdot & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdot & \cdot & \cdot & a_{m,2} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{1,n} & a_{2,n} & \cdot & \cdot & \cdot & a_{m,n} \end{bmatrix}^T$	Definition of transpose

$$= \begin{bmatrix} a_{1,1} & a_{1,2} & \cdot & \cdot & \cdot & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdot & \cdot & \cdot & a_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{m,1} & a_{m,2} & \cdot & \cdot & \cdot & a_{m,n} \end{bmatrix} \qquad \text{Definition of transpose}$$

$$= \quad A \qquad \text{Definition of A}$$

Since A was arbitrary, we have for all $A \in \mathbf{K}^{mxn}$ that $(A^T)^T = A$ as was to be proved.

Q.E.D.

Thus, if you take the transpose twice you get the original matrix back. This means that the function or mapping defined by the operation of taking the transpose is in some sense, its own inverse operation. (We call such functions or mappings **operators**.) If $f: \mathbf{K}^{m\times n} \to \mathbf{K}^{m\times n}$ is defined by $f(A) = A^T$, then $f^{-1}(B) = B^T$. However, $f \neq f^{-1}$ unless m=n. That is, taking the transpose of a square matrix is its own inverse function.

<u>DEFINITION #2.</u> If $A \in \mathbf{K}^{n\times n}$ (i.e. A is square) and $A = A^T$, then A is **symmetric**.

<u>COMPLEX CONJUGATE</u>.

<u>DEFINITION #3</u>. If $\underset{mxn}{A} = [\, a_{ij} \,] \in \mathbf{C}^{mxn}$, then $\underset{mxn}{\overline{A}} = [\, \overline{a}_{ij} \,] \in \mathbf{C}^{mxn}$. The **(complex) conjugate function** maps (or transforms) a matrix in $\mathbf{C}^{mxn}$ to (or into) a matrix in $\mathbf{C}^{mxn}$. Thus for the complex conjugate function C we have C: $\mathbf{C}^{mxn} \to \mathbf{C}^{mxn}$.

That is, the entry in the i^{th} row and j^{th} column is $\overline{a}_{ij}$ for 1,...,n, j=1,...,m. (i.e. the **complex conjugate** of A is obtained by taking the <u>complex conjugate</u> componentwise; that is, by taking the <u>complex conjugate</u> of each entry in the matrix.) Again we have not just defined one function, but rather an infinite number of functions, one for each set $\mathbf{K}^{mxn}$. We have named the function C, but we use $\overline{A}$ instead of C(A) for the element in the codomain $\mathbf{K}^{nxm}$ to which A is mapped. Note that unlike A^T there is no problem in defining $\overline{A}$ directly.

<u>PROPERTIES</u>. Again, there are lots of properties but most of them involve other operations not yet discussed. We give two at this time.

<u>THEOREM #2</u>. For all $A \in \mathbf{C}^{m\times n}$ = the set of all m×n complex valued matrices, $\overline{\overline{A}} = A$.

Thus if you take the complex conjugate twice you get the original matrix back. As with the transpose, taking the complex conjugate of a square matrix is its own inverse function.

THEOREM #3. $(\overline{A})^T = \overline{A^T}$.

That is, the operations of computing the complex conjugate and computing the transpose in some sense commute. That is, it does not matter which operation you do first. If A is square, then these two functions on $\mathbf{C}^{n\times n}$ do indeed commute in the technical sense of the word.

ADJOINT.

DEFINITION #4. If $\underset{m\times n}{A} = [\,a_{ij}\,] \in \mathbf{C}^{m\times n}$, then $\underset{m\times n}{A}{}^{*} = (\overline{A})^T \in \mathbf{C}^{n\times m}$. The **adjoint function** maps (or transforms) a matrix in $\mathbf{C}^{m\times n}$ to (or into) a matrix in $\mathbf{K}^{n\times m}$. Thus for the adjoint function *A*, we have , *A*: $\mathbf{C}^{m\times n} \rightarrow \mathbf{C}^{n\times m}$.

That is, the entry in the i^{th} row and j^{th} column is $\bar{a}_{ji}$ for 1,...,n, j=1,...,m. (i.e. the **adjoint** of A is obtained by taking the complex conjugate componentwise and then taking the transpose.) Again we have not just defined one function, but rather an infinite number of functions, one for each set $\mathbf{C}^{m\times n}$. We have named the function *A* (using italics to prevent confusion with the matrix A, but we use A* instead of *A*(A) for the element in the codomain $\mathbf{C}^{n\times m}$ to which A is mapped. . Note that having defined A^T and $\overline{A}$, A^* can be defined in terms of these two functions without having to directly refer to the entries of A.

PROPERTIES. Again, there are lots of properties but most of them involve other operations. Hence we give only one at this time.

THEOREM #4. $A^* = \overline{A^T}$.

That is, in computing A*, it does not matter whether you take the transpose or the complex conjugate first.

DEFINITION #4. If $A^* = A$, then A is said to be a **Hermitian matrix**.

THEOREM #5. If $A \in \mathbf{R}^{n\times n}$, then A is **symmetric** if and only if A is Hermitian.

Hermitian matrices (and hence real symmetric matrices) have nice properties; but more background is necessary before these can be discussed. Later, when we think of a matrix as (defining) an operator, then a Hermitian matrix is (defines) a self adjoint operator.

EXERCISES on Some Unary Matrix Operations

EXERCISE # 1. Compute the transpose of each of the following matrices

a) $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ b) A = [1, ½] c) $A = \begin{bmatrix} 1 & i & 1+i \\ 2 & 2+i & 0 \end{bmatrix}$ d) $A = \begin{bmatrix} \pi \\ e \\ \sqrt{2} \end{bmatrix}$

e) $A=[\pi,2]$ f) $A = \begin{bmatrix} i \\ 2+i \\ 3 \end{bmatrix}$ g) $A = \begin{bmatrix} 2 & 2i \\ 0 & -1 \end{bmatrix}$ h) $\begin{bmatrix} 0 & 0 \\ 1+i & 2+i \\ 1 & 0 \end{bmatrix}$ i) $A = [i,e,\pi]$

EXERCISE # 2. Compute the (complex) conjugate of each of the matrices in Exercise 1 above.

EXERCISE # 3. Compute the adjoint of each of the matrices in Exercise 1 above.

EXERCISE # 4. The notation $B = [b_{i\,j}]$ defines the entries of an arbitrary m×n matrix B. Write out this matrix as an array using a nine component representation (i.e, as is done in Equation (1). Now write out the matrix $C = [c_{ij}]$ as an array. Now write out the matrix C^T as an array using the nine component method.

EXERCISE # 5. Write a proof of Theorem #2. Since it is an identity, you can use the standard form for writing proofs of identities. Begin by defining the arbitrary matrix A. Represent it using the nine component method. Then start with the left hand side and go to the right hand side, justifying every step in a STATEMENT/REASON format.

EXERCISE # 6. Write a proof of Theorem #3. Since it is an identity, you can use the standard form for writing proofs of identities. Begin by defining the arbitrary matrix A. Represent it using the nine component method. Then start with the left hand side and go to the right hand side, justifying every step.

DEFINITION #1. Let A and B be $m \times n$ matrices in $\mathbf{K}^{m\times n}$ defined by

$$\underset{mxn}{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdot & \cdot & \cdot & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdot & \cdot & \cdot & a_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{m,1} & a_{m,1} & \cdot & \cdot & \cdot & a_{m,n} \end{bmatrix} = [a_{ij}], \quad \underset{mxn}{B} = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdot & \cdot & \cdot & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdot & \cdot & \cdot & b_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ b_{m,1} & b_{m,1} & \cdot & \cdot & \cdot & b_{m,n} \end{bmatrix} = [b_{ij}] \qquad (1)$$

We define the **sum** of A and B by

$$\underset{mxn}{A} + \underset{mxn}{B} = \begin{bmatrix} a_{1,1}+b_{1,1} & a_{1,2}+b_{1,2} & \cdot & \cdot & \cdot & a_{1,n}+b_{1,n} \\ a_{2,1}+b_{2,1} & a_{2,2}+b_{2,2} & \cdot & \cdot & \cdot & a_{2,n}+b_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{m,1}+b_{m,1} & a_{m,1}+b_{m,1} & \cdot & \cdot & \cdot & a_{m,n}+b_{m,n} \end{bmatrix} \in \mathbf{K}^{mxn} \qquad (2)$$

that is, $A + B \equiv C \equiv [c_{i\,j}]$ where $c_{i\,j} = a_{i\,j} + b_{i\,j}$ for $i = 1,..., m$ and $j = 1,..., n$.

We say that we add A and B componentwise. Note that it takes less space (but is not as graphic) to define the $i,j^{\underline{th}}$ component (i.e., $c_{i\,j}$) of $A + B = C$ then to write out the sum in rows and columns. Since m and n are arbitrary, in writing out the entries, we wrote only nine entries. However, these nine entries may give a more graphic depiction of a concept that helps with the visualization of matrix computations. 3×3 matrices are often used to represent physical quantities called **tensors**. Similar to referring to elements in $\mathbf{K}^n$, $\mathbf{K}^{n\times 1}$, and $\mathbf{K}^{1\times n}$ as "vectors" even when n>3 and $\mathbf{K} \neq \mathbf{R}$, we will refer to $A = a_{i\,j}$ (without parentheses or brackets) as **tensor notation** for any matrix A even when A is not 3×3. Thus $c_{i\,j} = a_{i\,j} + b_{i\,j}$ defines the matrix $c_{i\,j}$ as the sum of the matrices $a_{i\,j}$ and $b_{i\,j}$ using tensor notation. (The **Einstein summation convention** is discussed later, but it is **not assumed** unless so stated .)

As a mapping, + maps an element in $\mathbf{K}^{m\times n} \times \mathbf{K}^{m\times n}$ to an element in $\mathbf{K}^{m\times n}$, $+ : \mathbf{K}^{m\times n} \times \mathbf{K}^{m\times n} \to \mathbf{K}^{m\times n}$ where $\mathbf{K}^{m\times n} \times \mathbf{K}^{m\times n} = \{(A,B):A,B\in\mathbf{K}^{m\times n}\}$ is the **cross product** of $\mathbf{K}^{m\times n}$ with itself. The elements in $\mathbf{K}^{m\times n} \times \mathbf{K}^{m\times n}$ are **ordered pairs** of elements in $\mathbf{K}^{m\times n}$. Hence the order in which we add matrices could be important. Part b of Theorem 1 below establishes that it is not, so that matrix addition is **commutative**. Later we will see that multiplication of square matrices is a binary operation that is not commutative so that the order in which we multiply matrices is important.

PROPERTIES. Since the entries in a matrix are elements of a field, the field properties can be used to prove corresponding properties of matrix addition. These properties are then given the same names as the corresponding properties for fields.

THEOREM #1. Let A, B, and C be $m \times n$ matrices. Then

a. $\underset{m\times n}{A} + (\underset{m\times n}{B} + \underset{m\times n}{C}) = (\underset{m\times n}{A} + \underset{m\times n}{B}) + \underset{m\times n}{C}$ (matrix addition is **associative**) (3)

b. $\underset{m\times n}{A} + \underset{m\times n}{B} = \underset{m\times n}{B} + \underset{m\times n}{A}$ (matrix addition is **commutative**) (4)

THEOREM #2. There exists an $m \times n$ matrix **O** such that for all $m \times n$ matrices $\underset{m\times n}{A} + \underset{m\times n}{\mathbf{O}} = \underset{m\times n}{A}$ (existence of a **right additive identity** in $\mathbf{K}^{m\times n}$).

We call the right additive identity matrix the **zero matrix** since it can be shown to be the $m \times n$ matrix whose entries are all the zero element in the field **K**. (You may think of this element as the number zero in **R** or **C**.) To distinguish the zero matrix from the number zero we have denoted it by **O**; that is

$$\underset{m\times n}{\mathbf{O}} = \begin{bmatrix} 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & 0 \end{bmatrix} = [\,0\,] \in \mathbf{K}^{m\times n} \qquad (1)$$

THEOREM #3. The matrix $\mathbf{O} \in \mathbf{K}^{m\times n}$ is the unique **additive identity element** in $\mathbf{K}^{m\times n}$.

PARTIAL PROOF. The definition of and proof that **O** is an additive identity element is left to the exercises. We prove only that it is unique. To show that **O** is the unique element in $\mathbf{K}^{m\times n}$ such that for all $A \in \mathbf{K}^{m\times n}$, $A + \mathbf{O} = \mathbf{O} + A = A$, we assume that there is some other matrix, say $\Theta \in \mathbf{K}^{m\times n}$ that has the same property, that is, for all $A \in \mathbf{K}^{m\times n}$, $A + \Theta = \Theta + A = A$, and show that $\Theta = \mathbf{O}$. A proof can be written in several ways including showing that each element in Θ is the zero element of the field so that by the definition of equality of matrices we would have $\Theta = \mathbf{O}$. However, we can write a proof without having to consider the entries in Θ directly. Such a prove is considered to be more elegant. Although the conclusion, $\Theta = \mathbf{O}$ is an equality, it is not an identity. We will write the proof in the STATEMENT/REASON format, but not by starting with the left hand side. We start with a known identity.

STATEMENT	REASON
$\forall A \in \mathbf{K}^{mxn}$, $A = \mathbf{O} + A$	It is assumed that $\mathbf{O}$ has been established as <u>an</u> <u>additive identity element</u> for $\mathbf{K}^{mxn}$.
$\Theta = \mathbf{O} + \Theta$	Let $A = \Theta$ in this known identity.
$= \mathbf{O}$	Θ is assumed to be an identity so that for all $A \in \mathbf{K}^{mxn}$, $A + \Theta = \Theta + A = A$. This includes $A = \mathbf{O}$ so that $\mathbf{O} + \Theta = \mathbf{O}$.

Since we have shown that the assumed identity Θ must in fact be $\mathbf{O}$, (i.e., $\Theta = \mathbf{O}$), we have that $\mathbf{O}$ is the unique additive identity element for $\mathbf{K}^{mxn}$.

Q.E.D.

<u>THEOREM #4.</u> For every matrix $\underset{mxn}{A}$, there exists a matrix, call it $\underset{mxn}{B}$ such that

$$\underset{mxn}{A} + \underset{mxn}{B} = \underset{mxn}{\mathbf{O}} \qquad \text{(existence of a } \textbf{right additive inverse}\text{).}$$

It is easily shown that B is the matrix containing the negatives of the entries in A. That is, if $A = [a_{ij}]$ and $B = [b_{ij}]$, then $b_{ij} = -a_{ij}$. Hence we denote B by $-A$. (This use of the minus sign is technically different from the multiplication of a matrix by the scalar) 1. Multiplication of a matrix by a scalar is defined later. Only then can we <u>prove</u> that these are the same.)

<u>COROLLARY #5.</u> For every matrix $\underset{mxn}{A}$ the unique right additive inverse matrix, call it $-\underset{mxn}{A}$, is

$$-\underset{mxn}{A} = [-a_{ij}] \qquad \text{(Computation of the \underline{unique right additive inverse}).}$$

<u>THEOREM #6.</u> For every matrix $\underset{mxn}{A}$, the matrix $-\underset{mxn}{A}$ has the properties

$$\underset{mxn}{A} + (-\underset{mxn}{A}) = -\underset{mxn}{A} + \underset{mxn}{A} = \underset{mxn}{\mathbf{O}}. \qquad \text{(-A is a (right and left) } \textbf{additive inverse}\text{).}$$

Furthermore, this right and left inverse is unique.

<u>INVERSE OPERATION</u>. We define <u>subtraction</u> of matrices by
A - B = A + (-B). Thus to compute A) B, we first find the additive inverse matrix for B (i.e., C =) B where if $C = [c_{ij}]$ and $B = [b_{ij}]$, then c_{ij} =) b_{ij}). Then we <u>add</u> A to C =) B. Computationally, if $A = [a_{ij}]$ and $B = [b_{ij}]$ then

$$\underset{m\mathrm{x}n}{A} - \underset{m\mathrm{x}n}{B} = \begin{bmatrix} a_{1,1} - b_{1,1} & a_{1,2} - b_{1,2} & \cdot & \cdot & \cdot & a_{1,n} - b_{1,n} \\ a_{2,1} - b_{2,1} & a_{2,2} - b_{2,2} & \cdot & \cdot & \cdot & a_{2,n} - b_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{m,1} - b_{m,1} & a_{m,1} - b_{m,1} & \cdot & \cdot & \cdot & a_{m,n} - b_{m,n} \end{bmatrix} \in \mathbf{K}^{m\mathrm{x}n}$$

<u>THEOREM #7.</u> Let A, B $\in \mathbf{K}^{m\times n}$. Then $(A + B)^T = A^T + B^T$.

<u>THEOREM #8.</u> Let A, B $\in \mathbf{C}^{m\times n}$. Then the following hold:

1) $(\overline{A+B}) = \overline{A} + \overline{B}$

2) $(A + B)^* = A^* + B^*$

EXERCISES on Matrix Addition

<u>EXERCISE #1</u>. If possible, add A to B (i.e., find the sum A + B) if

a) $A = \begin{bmatrix} 1 & 1+i \\ 2 & 2-i \end{bmatrix}$, $B = \begin{bmatrix} 0 & 2i \\ 3 & 1-i \end{bmatrix}$ b) A = [1,2] , B = [1,2,3]

c) $A = \begin{bmatrix} 0 & 1+i \\ 2e & 3 \\ 1-i & i \end{bmatrix}$, $B = \begin{bmatrix} 1+i & 3 \\ 2 & 0 \\ 0 & 1-i \end{bmatrix}$ d) $A = \begin{bmatrix} 2+\sqrt{2}i \\ 0 \\ 1-\pi i \end{bmatrix}$, $B = \begin{bmatrix} 2+2\sqrt{2}i \\ 0 \\ 1-2\pi i \end{bmatrix}$

<u>EXERCISE #2</u>. Compute A+(B+C) (that is, first add B to C and then add A to the sum obtained) and (A+B)+C (that is, add A+B and then add this sum to C) and show that you get the same answer if

a) $A = \begin{bmatrix} 1 & 1+i \\ 2 & 2-i \end{bmatrix}$, $B = \begin{bmatrix} 0 & 2i \\ 3 & 1-i \end{bmatrix}$, $C = \begin{bmatrix} 2 & 1+i \\ 1 & 1-i \end{bmatrix}$ b) $A = \begin{bmatrix} 0 & 1+i \\ 2e & 3 \\ 1-i & i \end{bmatrix}$, $B = \begin{bmatrix} 1+i & 3 \\ 2 & 0 \\ 0 & 1-i \end{bmatrix}$ $C = \begin{bmatrix} 2 & 2+i \\ 3e & 3 \\ 3-i & i \end{bmatrix}$,

$A = \begin{bmatrix} 2+\sqrt{2}i \\ 0 \\ 1-\pi i \end{bmatrix}$, $B = \begin{bmatrix} 2+2\sqrt{2}i \\ 0 \\ 1-2\pi i \end{bmatrix}$ $C = \begin{bmatrix} 1+\sqrt{2}i \\ 0 \\ 2-\pi i \end{bmatrix}$

<u>EXERCISE #3</u>. Compute A+B (that is, add A to B) and B+A (that is, add B to A) and show that you get the same answer if:

a) $A = \begin{bmatrix} 1 & 1+i \\ 2 & 2-i \end{bmatrix}$, $B = \begin{bmatrix} 0 & 2i \\ 3 & 1-i \end{bmatrix}$ b) $A = \begin{bmatrix} 0 & 1+i \\ 2e & 3 \\ 1-i & i \end{bmatrix}$, $B = \begin{bmatrix} 1+i & 3 \\ 2 & 0 \\ 0 & 1-i \end{bmatrix}$ $A = \begin{bmatrix} 2+\sqrt{2}i \\ 0 \\ 1-\pi i \end{bmatrix}$, $B = \begin{bmatrix} 2+2\sqrt{2}i \\ 0 \\ 1-2\pi i \end{bmatrix}$

<u>EXERCISE #4</u>. Can you explain in one sentence why both commutativity and associativity hold for matrix addition? (Hint: They follow because of the corresponding properties for _________.)

(Fill in the blank)

Now elaborate.

EXERCISE #5. Find the additive inverse of A if

a) $\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 3 & -2 \end{bmatrix}$; b) $\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ c) $A = [i,e,\pi]$

EXERCISE # 6. Write a proof of Theorem #1. Since these are identities, use the standard form for writing proofs of identities. Begin by defining arbitrary matrices A and B. Represent them using at least the four corner entries. Then start with the left hand side and go to the right hand side, justifying every step.

EXERCISE # 7. Write a proof of Theorem #2. Since it is an identity, you can use the standard form for writing proofs of identities. Begin by defining the arbitrary matrix A. Represent it using at least the four corner entries. Then start with the left hand side and go to the right hand side, justifying every step.

EXERCISE # 8. Finish the proof of Theorem #3. Theorem #2 claims that **O** is a right additive identy, i.e., $\forall A \in \mathbf{K}^{mxn}$ we have the identity A+**O** = A. Thus we can use the standard form for writing proofs of identities to show that O is also a left additive identity. Begin by defining the arbitrary matrix A. Represent it using at least the four corner entries. Then define the additive inverse element **O** using at least the four corners. Then start with the left hand side and go to the right hand side, justifying every step. Hence **O** is an additive identity. We have shown uniqueness so that **O** is the additive identity element.

EXERCISE # 9. Write a proof of Theorem #4. Since it is an identity, you can use the standard form for writing proofs of identities. Begin by defining the arbitrary matrix A. Represent it using at least the four corner entries. Then define the additive inverse element $B = -A = [-a_{ij}]$ using at least the four corners. Then start with the left hand side and go to the right hand side, justifying every step.

EXERCISE # 10. Write a proof of Corollary #5. You can use the standard form for writing proofs of identities to show that −A is a right additive inverse. Then show uniqueness.

EXERCISE # 11. Write a proof of Theorem #6. Since it is an identity, you can use the standard form for writing proofs of identities.

EXERCISE # 12. Write a proof of Theorem #7. Since it is an identity, you can use the standard form for writing proofs of identities.

EXERCISE # 13. Write a proof of Theorem #8. Since these are identities, you can use the standard form for writing proofs of identities.

Componentwise multiplication is **not** what is usually called matrix multiplication. Although not usually developed in elementary matrix theory, it does not deserve to be ignored.

DEFINITION #1. Let A and B be m × n matrices in $\mathbf{K}^{m\times n}$ defined by

$$\underset{mxn}{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdot & \cdot & \cdot & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdot & \cdot & \cdot & a_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{m,1} & a_{m,1} & \cdot & \cdot & \cdot & a_{m,n} \end{bmatrix} = [a_{ij}], \quad \underset{mxn}{B} = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdot & \cdot & \cdot & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdot & \cdot & \cdot & b_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ b_{m,1} & b_{m,1} & \cdot & \cdot & \cdot & b_{m,n} \end{bmatrix} = [b_{ij}] \qquad (1)$$

Then we define the **componentwise product** of A and B by

$$\underset{mxn}{A} \otimes \underset{mxn}{B} = \begin{bmatrix} a_{1,1}b_{1,1} & a_{1,2}b_{1,2} & \cdot & \cdot & \cdot & a_{1,n}b_{1,n} \\ a_{2,1}b_{2,1} & a_{2,2}b_{2,2} & \cdot & \cdot & \cdot & a_{2,n}b_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{m,1}b_{m,1} & a_{m,1}b_{m,1} & \cdot & \cdot & \cdot & a_{m,n}b_{m,n} \end{bmatrix} \in \mathbf{K}^{mxn} \qquad (2)$$

that is, $A \otimes B \equiv C \equiv [c_{ij}]$ where $c_{ij} = a_{ij}\, b_{ij}$ for $i = 1,\ldots, m$ and $j = 1,\ldots, n$.

We say that we multiply A and B <u>componentwise</u>. Note that it takes less space (but is not as graphic) to define the $i,j^{\underline{th}}$ component (i.e., $c_{i\,j}$) of $C = A \otimes B$ then to write out the product in rows and columns. Even in writing out the entries we were terse since we wrote only nine entries of the array. Again, since this nine element technique often gives a more graphic depiction of a concept, it will often be used to help with the visualization of matrix computations.

<u>PROPERTIES</u>. Since the entries in a matrix are elements of a field, the field properties can be used to prove corresponding properties for componentwise matrix multiplication.

<u>THEOREM #1</u>. Let A, B, and C be m × n matrices. Then

a. $\underset{mxn}{A} \otimes (\underset{mxn}{B} \otimes \underset{mxn}{C}) = (\underset{mxn}{A} \otimes \underset{mxn}{B}) \otimes \underset{mxn}{C}$ (componentwise matrix mult. is associative) (3)

b. $\underset{mxn}{A} \otimes \underset{mxn}{B} = \underset{mxn}{B} \otimes \underset{mxn}{A}$ (componentwise matrix mult. is commutative) (4)

THEOREM #2. There exists a unique $m \times n$ matrix **1** such that for all $m \times n$ matrices

$$\underset{m\text{x}n}{A} \otimes \underset{m\text{x}n}{\mathbf{1}} = \underset{m\text{x}n}{A} \otimes \underset{m\text{x}n}{\mathbf{1}} = \underset{m\text{x}n}{A}$$ (existence of a **componentwise multiplicative identity**).

We call the componentwise multiplicative identity matrix the **one matrix** since it can be shown to be the $m \times n$ matrix whose entries are all ones. To distinguish the one matrix from the number one we have denoted it by **1**; that is

$$\underset{m\text{x}n}{\mathbf{1}} = \begin{bmatrix} 1 & 1 & \cdot & \cdot & \cdot & 1 \\ 1 & 1 & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ 1 & 1 & \cdot & \cdot & \cdot & 1 \end{bmatrix} = [\,1\,] \in \mathbf{K}^{m\text{x}n} \qquad (5)$$

THEOREM #3. For every matrix $\underset{m\text{x}n}{A} = [a_{ij}]$ such that $a_{ij} \neq 0$ for i =1, 2, ... , m, and j = 1, 2, ... , n, there exists a unique matrix, call it $\underset{m\text{x}n}{B}$ such that

$$\underset{m\text{x}n}{A} \otimes \underset{m\text{x}n}{B} = \underset{m\text{x}n}{B} \otimes \underset{m\text{x}n}{A} = \underset{m\text{x}n}{\mathbf{1}}$$ (existence of a unique **componentwise multiplicative inverse**).

It is easily shown that B is the matrix containing the multiplicative inverses of the entries in A. That is, if $A = [a_{ij}]$ and $B = [b_{ij}]$, then $b_{ij} = 1/a_{ij}$. Hence we denote B by 1/A.

INVERSE OPERATION. We define componentwise division of matrices by
$A / B = A \otimes (1/B)$. Thus to compute A / B, we first find the componentwise multiplicative inverse matrix for B (i.e., C = 1/B where if $C = [c_{ij}]$ and $B = [b_{ij}]$, then $c_{ij} = 1/b_{ij}$). Then we componentwise multiply A by C = 1/B. Computationally, if $A = [a_{ij}]$ and $B = [b_{ij}]$ then

$$\underset{m\text{x}n}{A} / \underset{m\text{x}n}{B} = \begin{bmatrix} a_{1,1}/b_{1,1} & a_{1,2}/b_{1,2} & \cdot & \cdot & \cdot & a_{1,n}/b_{1,n} \\ a_{2,1}/b_{2,1} & a_{2,2}/b_{2,2} & \cdot & \cdot & \cdot & a_{2,n}/b_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{m,1}/b_{m,1} & a_{m,1}/b_{m,1} & \cdot & \cdot & \cdot & a_{m,n}/b_{m,n} \end{bmatrix} \in \mathbf{K}^{m\text{x}n} \qquad (6)$$

EXERCISES on Componentwise Matrix Multiplication

EXERCISE #1. If possible, multiply componentwise A to B (i.e., find the componentwise product A $\circ_{cm}$ B) if

a) $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$ b) A = [1,2] , B = [1,2,3]

c) $\mathbf{A} = \begin{bmatrix} 1 & 0 & -1 \\ 2 & -1 & 3 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 1 & -3 \end{bmatrix}$ d) $\mathbf{A} = \begin{bmatrix} \pi \\ e \\ \sqrt{2} \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 1 + \pi \\ 2e \\ 3 + \sqrt{2} \end{bmatrix}$

EXERCISE #2. Compute A $\circ_{cm}$ (B $\circ_{cm}$ C) (that is, first componentwise multiply B and C and then componentwise multiply A and the product obtained) and (A $\circ_{cm}$ B) $\circ_{cm}$ C (that is, first componentwise multiply A and B and then componentwise multiply the product obtained with C) and show that you get the same answer if

a) $\mathbf{A} = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 5 & -3 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 2 & 5 & 1 \\ 0 & -5 & 2 \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix}$

b) $\mathbf{A} = [\pi, 2]$, $\mathbf{B} = \left[1 + 3\sqrt{2}, \sqrt{2}\right]$, $\mathbf{C} = \left[0, 1 + \sqrt{2}\right]$

EXERCISE #3. Compute A$\otimes$B and B$\otimes$A and show that you get the same answer if:

a) $\mathbf{A} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ b) $\mathbf{A} = \begin{bmatrix} 2 & 2i \\ 0 & -1 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 0 & i \\ 0 & 0 \end{bmatrix}$

EXERCISE #4. Can you explain in one sentence why commutativity and associativity hold for componentwise matrix multiplication? (Hint: They follow because of the corresponding properties for ____.)
(Fill in the blank, then elaborate)

EXERCISE #5. If possible, find the componentwise multiplicative inverse of A if

a) $\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 3 & -2 \end{bmatrix}$; b) $\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ c) A = [i,e,π]

EXERCISE # 6. Write a proof of Theorem #1. Since these are identities, use the standard form for writing proofs of identities. Begin by defining arbitrary matrices A and B. Represent them using at least the four corner entries. Then start with the left hand side and go to the right hand side, justifying every step.

EXERCISE # 7. Write a proof of Theorem #2. Since it is an identity, you can use the standard form for writing proofs of identities. Begin by defining the arbitrary matrix A. Represent it using at least the four corner entries. Then start with the left hand side and go to the right hand side, justifying every step.

EXERCISE # 8. Write a proof of Theorem #3. Since it is an identity, you can use the standard form for writing proofs of identities. Begin by defining the arbitrary matrix A. Represent it using at least the four corner entries. Then define the additive inverse element $B = -A = [-a_{ij}]$ using at least the four corners. Then start with the left hand side and go to the right hand side, justifying every step.

DEFINITION #1. Let $\underset{m\times n}{A}$ be an m×n matrix and α a scalar ($\alpha \in \mathbf{K}$ where **K** is a field):

$$\underset{mxn}{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdot & \cdot & \cdot & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdot & \cdot & \cdot & a_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{m,1} & a_{m,1} & \cdot & \cdot & \cdot & a_{m,n} \end{bmatrix} = [a_{ij}] \in \mathbf{K}^{mxn}, \qquad \alpha\in \mathbf{K} \text{ (e.g., } \mathbf{K} = \mathbf{R} \text{ or } \mathbf{K} = \mathbf{C}) \tag{1}$$

then we define the product of α and A (called **multiplication by a scalar**, but not scalar product) by

$$\alpha \underset{mxn}{A} = \begin{bmatrix} \alpha a_{1,1} & \alpha a_{1,2} & \cdot & \cdot & \cdot & \alpha a_{1,n} \\ \alpha a_{2,1} & \alpha a_{2,2} & \cdot & \cdot & \cdot & \alpha a_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \alpha a_{m,1} & \alpha a_{m,1} & \cdot & \cdot & \cdot & \alpha a_{m,n} \end{bmatrix} = [\ \alpha\, a_{ij}\ ] \in \mathbf{K}^{mxn} \tag{1}$$

that is, $\alpha A \equiv C = [c_{ij}]$ where $c_{ij} = \alpha a_{ij}$ for i = 1,..., m and j = 1,..., m. That is, we multiply each component in A by α. Again defining an arbitrary component c_{ij} of αA takes less space but is less graphic, that is, does not give a good visualization of the operation. However, you should learn to provide that visualization, (i.e., look at $c_{ij} = \alpha a_{ij}$ and visualize each component being multiplied by α, for example, the nine element graphic above or the four corner graphic.) Sometimes we place the scalar on the right hand side (RHS): $\alpha A = \alpha[a_{ij}] = [a_{ij}]\alpha = A\alpha$.

PROPERTIES. The following theorems can be proved.

THEOREM #1. Let A be an m×n matrix and 1 be the multiplicative identity in the associated scalar field **K** (e.g., $1 \in \mathbf{R}$ or $1 \in \mathbf{C}$), then 1A = A.

THEOREM #2. Let A and B be an m×n matrix and α and β be scalars in **K**(e.g., $\alpha,\beta \in \mathbf{R}$ or $\alpha,\beta \in \mathbf{C}$), then

a. $\alpha(\beta A) = (\alpha\beta)A$ (Note the difference in the meaning of the two sets of parentheses.)
b. $(\alpha + \beta)A = (\alpha A) + (\beta A)$ (Note the difference in the meaning of the two plus signs.)
c. $\alpha(A+B) = (\alpha A) + (\beta B)$ (What about parentheses and plus signs here?)

INVERSE OPERATION. If α is a nonzero scalar, the division of a matrix A by the scalar α is defined to be multiplication by the multiplicative inverse of the scalar. Thus $A/\alpha = (1/\alpha)\,A$.

THEOREM #3. Let $A, B \in \mathbf{C}^{m\times n}$. Then the following hold:

1) $(\alpha A)^T = \alpha A^T$

2) $\overline{\alpha A} = \bar{\alpha}\bar{A}$

3) $(\alpha A)^* = \bar{\alpha}\, A^*$

EXERCISES on Multiplication by a Scalar

EXERCISE #1. Multiply A by the scalar α if

a) $\alpha = 2$ and $A = \begin{bmatrix} 1 & 1+i \\ 2 & 2-i \end{bmatrix}$ b) $\alpha = 2+i$ and $A = [1,2]$ c) $\alpha = 1$ and $A = \begin{bmatrix} 0 & 1+i \\ 2e & 3 \\ 1-i & i \end{bmatrix}$

d) $\alpha = 1/2$ and $A = \begin{bmatrix} 2+\sqrt{2}i \\ 0 \\ 1-\pi i \end{bmatrix}$

EXERCISE #2. Compute $(\alpha A)^T$ (that is, first multiply α by A and take the transpose) and then αA^T (that is, multiply α by A transpose) and show that you get the same answer if

a) $\alpha = 2$ and $A = \begin{bmatrix} 1 & 1+i \\ 2 & 2-i \end{bmatrix}$ b) $\alpha = 2+i$ and $A = [1,2]$ c) $\alpha = 1$ and $A = \begin{bmatrix} 0 & 1+i \\ 2e & 3 \\ 1-i & i \end{bmatrix}$

d) $\alpha = 1/2$ and $A = \begin{bmatrix} 2+\sqrt{2}i \\ 0 \\ 1-\pi i \end{bmatrix}$

EXERCISE #3. Compute $\overline{\alpha A}$ (that is, multiply α by A and then compute the complex conjugate) and then compute $\bar{\alpha}\bar{A}$ (that is, compute $\bar{\alpha}$ and then $\bar{A}$ and then multiply $\bar{\alpha}$ by $\bar{A}$) and then show that you get the same answer if:

a) $\alpha = 2$ and $A = \begin{bmatrix} 1 & 1+i \\ 2 & 2-i \end{bmatrix}$ b) $\alpha = 2+i$ and $A = [1,2]$ c) $\alpha = 1$ and $A = \begin{bmatrix} 0 & 1+i \\ 2e & 3 \\ 1-i & i \end{bmatrix}$

d) $\alpha = ½$ and $A = \begin{bmatrix} 2+\sqrt{2}i \\ 0 \\ 1-\pi i \end{bmatrix}$

DEFINITION #1. Let $\vec{x} = [x_1, x_2, ..., x_n]$ and $\vec{y} = [y_1, y_2, ..., y_n]$ be row vectors in $\mathbf{R}^{1\times n}$ or $\mathbf{C}^{1\times n}$ (we could just as well use column vectors). Then define the **scalar product** of $\vec{x}$ and $\vec{y}$ by

$$\vec{x} \circ \vec{y} = x_1\overline{y}_1 + x_2\overline{y}_2 + \ldots + x_n\overline{y}_n = \sum_{i=1}^{n} x_i\overline{y}_i$$

In $\mathbf{R}^2$ and $\mathbf{R}^3$, this is usually called the **dot product**. (In an abstract setting, this operation is usually called an **inner product**.)

PROPERTIES. The following theorem can be proved.

THEOREM #1. Let $\vec{x} = [x_1, x_2, ..., x_n]$, $\vec{y} = [y_1, y_2, ..., y_n]$, and $\vec{z} = [z_1, z_2, ..., z_n]$ be row vectors in $\mathbf{R}^{1\times n}$ or $\mathbf{C}^{1\times n}$; we could just as well use column vectors) and α be a scalar. Then

a. $\vec{x}\cdot\vec{y} = \overline{\vec{y}\cdot\vec{x}} = \overline{\vec{y}}\cdot\overline{\vec{x}}$
b. $(\vec{x}+\vec{y})\cdot\vec{z} = \vec{x}\cdot\vec{y}+\vec{x}\cdot\vec{z}$
c. $(\alpha\vec{x})\cdot\vec{y} = \alpha(\vec{x}\cdot\vec{y})$
d. $\vec{x}\cdot\vec{x} \geq 0$
$\vec{x}\cdot\vec{x} = 0$ if and only if $\vec{x} = \vec{0}$

The term **inner product** is used for an operation on an abstract vector space if it has all of the properties given in Theorem #1. Hence we have that the operation of scalar product defined above is an example of an inner product and hence $\mathbf{R}^{1\times n}$ and $\mathbf{K}^{1\times n}$ (and $\mathbf{R}^{n\times 1}$ and $\mathbf{C}^{n\times 1}$ and $\mathbf{R}^n$ and $\mathbf{C}^n$) are **inner product spaces**.

EXERCISES on Dot or Scalar Product.

EXERCISE #1. Compute $\vec{x}\cdot\vec{y}$ if
a) $\vec{x}=[1,2,1]$, $\vec{y}=[2,3,1]$ b) $\vec{x}=[1,2,3,4]$, $\vec{y}=[2,4,0,1]$ c) $\vec{x}=[1,0,1,0]$, $\vec{y}=[1,2,0,1]$

EXERCISE #2. Compute $\vec{x}\cdot\vec{y}$ if
a) $\vec{x}=[1+i,2+i,0,1], \vec{y}=[1+i,2+i,0,1]$ b) $\vec{x}=[1+i,0,0,i], \vec{y}=[1+i,2+i,0,1]$
c) $\vec{x}=[1,0,1,i], \vec{y}=[1+i,2+i,0,1]$

Matrix multiplication should probably be called matrix composition, but the term matrix multiplication is standard and we will use it.

DEFINITION #1. Let $\underset{m\times p}{A} = [\,a_{ij}\,]$, $\underset{p\times n}{B} = [\,a_{ij}\,]$. Then $\underset{m\times n}{C} = AB = [\,c_{ij}\,]$ where $c_{ij} = \sum_{i=1}^{p} a_{ik} b_{kj}$.

Although the definition of the product matrix C is given without reference to the scalar product, it is useful to note that the element c_{ij} is obtained by computing the scalar product of the i^{th} row of A with the j^{th} column of B and placing this result in the i^{th} row and j^{th} column of C. Using tensor notation, we may just write $c_{ij} = \sum_{i=1}^{p} a_{ik} b_{kj}$. If we adopt the **Einstein summation convention** of summing over repeated indices, then this may be written as $c_{ij} = a_{ik}\, b_{kj}$. This assumes that the values of m, n, and p are known.

EXAMPLE.

$$\underset{A}{\begin{bmatrix} -1 & 2 & 1 \\ 0 & 2 & 5 \\ 1 & 3 & 1 \end{bmatrix}} \underset{B}{\begin{bmatrix} 0 & 3 & 2 \\ 1 & -1 & 3 \\ 1 & 1 & 1 \end{bmatrix}} = \underset{C}{\begin{bmatrix} 3 & -4 & 5 \\ 7 & 3 & 11 \\ 4 & 1 & 12 \end{bmatrix}}$$

$c_{11} = [-1,2,1]\circ[0,1,1] = [-1,2,1][0,1,1]^T \quad = (-1)(0) + (2)(1) + (1)(1) = 3$

$c_{12} = [-1,2,1]\circ[3,-1,1] = [-1,2,1][3,-1,1]^T \quad = (-1)(3) + (2)(-1) + (1)(1) = -4$

.
.
.

SIZE REQUIREMENT: In order to take the dot product of the i^{th} row of A with the j^{th} column of B, they must have the same number of elements. Thus the number of columns of A must be the same as the number of rows of B.

$$\underset{n\times p}{A}\ \underset{p\times n}{B} = \underset{n\times m}{C}$$

THE SCALAR PRODUCT IN TERMS OF MATRIX MULTIPLICATION. The dot product of two row vectors in $\mathbf{R}^{1\times n}$ can be given in terms of matrix multiplication. Let $\vec{x} = [x_1,x_2,...,x_n]$ and $\vec{y} = [y_1,y_2,...,y_n]$ be row vectors in $\mathbf{R}^{1\times n}$. Then

$$\vec{x}\cdot\vec{y} = x_1y_1 + x_2y_2 + \ldots + x_ny_n = \sum_{i=1}^{n} x_iy_i = \underset{1xn}{\vec{x}}\ \underset{nx1}{\vec{y}}^{T}$$

If $\vec{x}$ and $\vec{y}$ are defined as column vectors, we have

$$\vec{x}\cdot\vec{y} = x_1y_1 + x_2y_2 + \ldots + x_ny_n = \sum_{i=1}^{n} x_iy_i = \underset{1xn}{\vec{x}}^{T}\ \underset{nx1}{\vec{y}}$$

Using tensor notation, including the Einstein summation convention, we have $\vec{x}\cdot\vec{y} = x_i\, y_i$.

<u>PROPERTIES</u>. For multiplication properties, we must first be sure that all of the operations are possible. Note that unless A and B are square, we can not compute both AB and BA. Even for square matrices, AB does not always equal BA (except for 1×1's). However, we do have:

<u>THEOREM #1.</u> Let $\underset{mxp}{A}$, $\underset{pxr}{B}$, and $\underset{rxn}{C}$ be matrices so that the multiplications $\underset{mxp}{A}\ \underset{pxr}{B}$, $\underset{pxr}{B}\ \underset{rxn}{C}$, $\underset{mxp}{A}\,(\underset{pxr}{B}\ \underset{rxn}{C})$, and $(\underset{mxp}{A}\ \underset{pxr}{B})\,\underset{rxn}{C}$ are all possible. Then

A(BC) = (AB)C. matrix multiplication is **associative**

<u>THEOREM #2.</u> Let $\underset{mxp}{A}$, $\underset{pxn}{B}$, and $\underset{pxn}{C}$ be matrices so that the multiplications $\underset{mxp}{A}\ \underset{pxn}{B}$, $\underset{mxp}{A}\ \underset{pxn}{C}$ and the additions $\underset{pxn}{B} + \underset{pxn}{C}$ and $\underset{mxp}{A}\ \underset{pxn}{B} + \underset{mxp}{A}\ \underset{pxn}{C}$ are all possible. Then

A(B + C) = AB + AC matrix multiplication on the left **distributes** over matric addition

Now let $\underset{mxp}{A}$, $\underset{mxp}{B}$, and $\underset{pxn}{C}$ be matrices so that the multiplications $\underset{mxp}{A}\ \underset{pxn}{B}\ \underset{pxn}{C}$, $\underset{mxp}{B}\ \underset{pxn}{C}$, and the additions $\underset{mxp}{A} + \underset{mxp}{B}$, and $\underset{mxp}{A}\ \underset{pxn}{C} + \underset{mxp}{B}\ \underset{pxn}{C}$ are all possible. Then

(A+B)C = AC + BC matrix multiplication on the right **distributes** over matric addition

<u>THEOREM #3.</u> Let $\underset{mxp}{A}$ and $\underset{pxn}{B}$ be matrices so that the matrix multiplications AB, (αA)B, and A(αB) are all possible. Then (αA)B = A(αB) = α(AB).

EXERCISES on Matrix Multiplication

EXERCISE #1. If possible, multiply A times B (i.e., find the product AB) if

a) $A = \begin{bmatrix} 1 & 1+i \\ 2 & 2-i \end{bmatrix}$, $B = \begin{bmatrix} 0 & 2i \\ 3 & 1-i \end{bmatrix}$ b) $A = \begin{bmatrix} 1 & 1+i \\ 1 & 1-i \end{bmatrix}$, $B = \begin{bmatrix} 0 & 2i \\ 3 & 1-i \end{bmatrix}$ c) A = [1,2] , B = [1,2,3]

d) $A = \begin{bmatrix} 0 & 1+i \\ 2e & 3 \\ 1-i & i \end{bmatrix}$, $B = \begin{bmatrix} 1+i & 3 \\ 2 & 0 \\ 0 & 1-i \end{bmatrix}$ e) $A = \begin{bmatrix} 1 & 0 \\ 2 & 2-i \end{bmatrix}$, $B = \begin{bmatrix} 0 & 2i \\ 3 & 1-i \end{bmatrix}$ f) $A = \begin{bmatrix} 1 & 1+i \\ 2 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 2i \\ 3 & 1-i \end{bmatrix}$

g) $A = \begin{bmatrix} 2+\sqrt{2}i \\ 0 \\ 1-\pi i \end{bmatrix}$, $B = \begin{bmatrix} 2+2\sqrt{2}i \\ 0 \\ 1-2\pi i \end{bmatrix}$ h) $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 0 \\ 1+i & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

EXERCISE #2. . If posible, compute A(B+C) (that is, first add B to C and then add A times this sum) and then AB and BC (that is, multiply AB and then BC) and then show that you get A(B+C) = AB +AC if

a) $A = \begin{bmatrix} 1 & 1+i \\ 2 & 2-i \end{bmatrix}$, $B = \begin{bmatrix} 0 & 2i \\ 3 & 1-i \end{bmatrix}$, $C = \begin{bmatrix} 2 & 1+i \\ 1 & 1-i \end{bmatrix}$ b) $A = \begin{bmatrix} 0 & 1+i \\ 2e & 3 \\ 1-i & i \end{bmatrix}$, $B = \begin{bmatrix} 1+i & 3 \\ 2 & 0 \\ 0 & 1-i \end{bmatrix}$ $C = \begin{bmatrix} 2 & 2+i \\ 3e & 3 \\ 3-i & i \end{bmatrix}$

If A and B are square, then we can compute both AB and BA. Unfortunately, these may not be the same.

<u>THEOREM #1.</u> If $n > 1$, then there exists $A, B \in \mathbf{R}^{n\times n}$ such that $AB \neq BA$. Thus matrix multiplication is **<u>not</u> commutative**.

Thus AB=BA is not an identity. Can you give a **counter example** for n=2? (i.e. an example where $AB \neq BA$. See Exercise #2.)

<u>DEFINITION #1</u>. For square matrices, there is a **multiplicative identity element**. We define the n×n matrix I by

$$\underset{m\mathrm{x}n}{I} = \begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & 1 \end{bmatrix}$$ One's down the diagonal. Zero's everywhere else.

<u>THEOREM #2</u>. We have $\underset{n\mathrm{x}n}{A}\ \underset{n\mathrm{x}n}{I} = \underset{n\mathrm{x}n}{I}\ \underset{n\mathrm{x}n}{A} = \underset{n\mathrm{x}n}{A} \qquad \forall A \in \mathbf{K}^{n\mathrm{x}n}$

<u>DEFINITION #2</u>. If there exists B such that AB = I., then B is a **right (multiplicative) inverse** of A. If there exists C such that CA = I., then C is a **left (multiplicative) inverse** of A. If AB = BA = I, then B is <u>a</u> (**multiplicative) inverse** of A and we say that A is **invertible**. If B is the only matrix with the property that AB = BA = I, then B is <u>the</u> **inverse** of A. If A has a unique inverse, then we say A is **nonsingular** and denote its inverse by A^{-1}.

<u>THEOREM #3</u>. Th identity matrix is its own inverse.

Later (Chapter 9) we show that if A has a right and a left inverse, then it has a unique inverse. Hence we prove that A is invertible if and only if it is nonsingular. Even later, we show that if A has a right (or left) inverse, then it has a unique inverse. Thus, even though matrix multiplication is not commutative, a right inverse is always a left inverse and is indeed <u>the</u> inverse. Some matrices have inverses; others do not. Unfortunately, it is usually not easy to look at a matrix and determine whether or not it has a (multiplicative) inverse.

<u>THEOREM #4</u>. There exist $A, B \in \mathbf{R}^{n\times n}$ such that $A \neq I$ is invertible and B has no inverse.

INVERSE OPERATION. If B has a right and left inverse then it is a unique inverse ((i.e., $\exists\, B^{-1}$ such that $B^{-1}B = BB^{-1} = I$) and we can define **Right Division** AB^{-1} and **Left Division** $B^{-1}A$ of A by B (provided B^{-1} exists). But since matrix multiplication is not commutative, we do not know that these are the same. Hence $\frac{A}{B}$ is not well defined since no indication of whether we mean left or right division is given.

EXERCISES on Matrix Algebra for Square Matrices

EXERCISE #1. True or False.

_____ 1. If A and B are square, then we can compute both AB and BA.
_____ 2. If n >1, then there exists A, $B \in \mathbf{R}^{n\times n}$ such that $AB \neq BA$.
_____ 3. Matrix multiplication is not commutative.
_____ 4. AB=BA is not an identity.
_____ 5. For square matrices, there is a multiplicative identity element, namely the n×n matrix I, given by

$$I_{nxn} = \begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & 1 \end{bmatrix}.$$

_____ 6. $\forall\, A \in \mathbf{K}^{nxn}$ we have $A_{nxn} I_{nxn} = I_{nxn} A_{nxn} = A_{nxn}$
_____ 7. If there exists B such that AB = I., then B is a right (multiplicative) inverse of A.

_____ 8. If there exists C such that CA = I., then C is a left (multiplicative) inverse of A.
_____ 9. If AB = BA = I, then B is a multiplicative inverse of A and we say that A is invertible.
_____ 10. If B is the only matrix with the property that AB = BA = I, then B is the inverse of A.
_____ 11. If A has a unique inverse, then we say A is nonsingular and denote its inverse by A^{-1}.
_____ 12. The identity matrix is its own inverse.
_____ 13. If A has a right and a left inverse, then it has a unique inverse.
_____ 14. A is invertible if and only if it is nonsingular.
_____ 15. If A has a right (or left) inverse, then it has a unique inverse.
_____ 16. Even though matrix multiplication is not commutative, a right inverse is always a left inverse.
_____ 17. The inverse of a matrix is unique.
_____ 18. Some matrices have inverses; others do not.
_____ 19. It is usually not easy to look at a matrix and determine whether or not it has a (multiplicative) inverse.
_____ 20. There exist A, $B \in \mathbf{R}^{n\times n}$ such that $A \neq I$ is invertible and B has no inverse

EXERCISE #2. Let $\alpha = 2$, $A = \begin{bmatrix} 1+i & 1-i \\ 1 & 0 \end{bmatrix}$, and $B = \begin{bmatrix} 1 & 0 \\ i & 1+i \end{bmatrix}$. Compute the following:

$\overline{A}$ = ____________. A^T =_____________. A^* = _____________. αA = _____________.

A+B =_____________. AB =______________.

EXERCISE #3. Let $\alpha = 3$, $A = \begin{bmatrix} i & 1-i \\ 0 & 1+i \end{bmatrix}$, and $B = \begin{bmatrix} 1 & 0 \\ i & 1+i \end{bmatrix}$. Compute the following:

$\overline{A}$ = ____________. A^T =_____________. A^* = _____________. αA = _____________.

A+B =_____________. AB =______________.

EXERCISE #4. Solve $\underset{2x2}{A}\underset{2x1}{\vec{x}} = \underset{2x1}{\vec{b}}$ where $A = \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} 1 \\ i \end{bmatrix}$.

EXERCISE #5. Solve $\underset{2x2}{A}\underset{2x1}{\vec{x}} = \underset{2x1}{\vec{b}}$ where $A = \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

EXERCISE #6 Solve $\underset{2x2}{A}\underset{2x1}{\vec{x}} = \underset{2x1}{\vec{b}}$ where $A = \begin{bmatrix} 1 & i \\ i & 0 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} 1 \\ i \end{bmatrix}$

Recall that if **K** is a field and $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then the following properties of matrices are true.

<u>THEOREM #1.</u> Let $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**. Then for all A, B, C $\in \mathbf{K}^{m\times n}$, we have

1) $A + (B + C) = (A + B) + C$ — associativity of matrix addition
2) $A + B = B + A$ — commutativity of matrix addition
3) There exists a unique matrix **O** such that for every matrix A ε $\mathbf{K}^{m\times n}$, $A + \mathbf{O} = A$.
4) For each A ε $\mathbf{K}^{m\times n}$, there exist a unique matrix called $-A$ such that $A + (-A) = \mathbf{O}$.

By using these properties, but without resorting to looking at the components of the matrices we can prove

<u>THEOREM #2.</u> Let $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**. Then for all A, B, C $\in \mathbf{K}^{m\times n}$, we have

1) $\mathbf{O} + A = A$.
2) $(-B) + (A + B) = (A + B) + (-B) = A$
3) $-(A + B) = (-A) + (-B)$. The additive inverse of a sum is the sum of the additive inverses.
4) If $A + B = A + C$, then $B = C$ This is the cancellation law (for addition).

The proofs of properties 1) to 3) in Theorem #2 are easily proved using the standard method for proving identities in STATEMENT/REASON format. However none of the reasons rely on looking at the components of the matrices and hence do not rely directly on the properties of the underlying field. Now note that property 4) in Theorem #2 is not an identity. The conclusion is an equality, but it is a **conditional equality**. Although one could write a somewhat contorted proof of the concluding equality (B = C) by starting with one side and using the **substitution axiom of equality** to achieve the other side, a better proof is achieved by simply adding the same element to both sides of the given equation. Explanations of why the element to be added exists (Property 4) in Theorem #4) and why you can add the same element to both sides of an equation and the result is the same (axiom of equality) are needed.

The properties given in Theorem #1 establish $\mathbf{K}^{m\times n}$ as a **Abelian** (or **commutative**) **group**. Since only these properties are needed to prove Theorem #2, we see that any mathematical structure having these properties (i.e., any commutative group) also has the properties given in Theorem #2. We refer to both the defining properties given in Theorem #1 and the resulting properties given in Theorem #2 as **Abelian group properties**.

Now recall that if **K** is a field and $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then the following additional properties of matrices have been proved.

<u>THEOREM #3.</u> Let $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**. Then for all scalars α ß $\in$ **K** and A, B $\in \mathbf{K}^{m\times n}$, we have

1) $\alpha (ß A) = (\alpha ß) A$ (Note that $(\alpha ß)$ indicates multiplication in the field **K**.)
2) $(\alpha + ß) A = \alpha A + ß A$
3) $\alpha (A + B) = \alpha A + \alpha B$
4) $1 A = A$.

The properties in Theorem #4 below are easy to prove directly for matrices. However, they can also be proved by using the properties in Theorem #1 and Theorem #3 (and the properties in Theorem #2 since these can be proved using only Theorem #1), but without resorting to looking at the components of the matrices.

THEOREM #4. Let $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**. Then for all scalars $\alpha \in \mathbf{K}$ and $A \in \mathbf{K}^{m\times n}$, we have

1) $0A = \mathbf{O}$,
2) $\alpha \mathbf{O} = \mathbf{O}$,
3) If $\alpha A = \mathbf{O}$, then either $\alpha = 0$ or $A = \mathbf{O}$.

The first two properties are identities and can be proved using the standard method. Some might argue that such proofs can become contorted and other methods may be clearer. The third is neither an identity nor a conditional equality. The hypothesis is an equality, but the conclusion states that there are only two possible reasons why the suppose equality could be true, either of the possibilities given in 1) or 2) (or both) but no others.

The properties given in Theorems #1 and #3 establish $\mathbf{K}^{m\times n}$ as a **vector space** (see Chapter 2-3) over **K**. Since only properties in Theorems #1 and #3 are needed to prove Theorems #2 and #4, we see that any mathematical structure having the properties given in Theorems #1 and #3 (i.e., any vector space) also has the properties given in Theorems #2 and 4. We refer to both the defining properties given in Theorems #1 and #3 and the resulting properties given in Theorems #2 and #4 as **vector space properties**. We consider additional vector spaces in the next chapter. Later, we consider more propertiess of (multiplicative) inverses of square matrices.

EXERCISES on Additional Matrix Properties

EXERCISE #1. True or False.

_____ 1.If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then for all A, B, C $\in\mathbf{K}^{m\times n}$, we have
$A + (B + C) = (A + B) + C$.

_____ 2.If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then for all A, B, C $\in\mathbf{K}^{m\times n}$, we have
the associativity of matrix addition.

_____ 3.If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then for all A, B $\in\mathbf{K}^{m\times n}$, we have
$A + B = B + A$.

_____ 4.If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then for all A, B $\in\mathbf{K}^{m\times n}$, we have
the commutativity of matrix addition.

_____ 5.If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then we have that there exists a unique matrix **O** such that for every matrix A $\varepsilon\,\mathbf{K}^{m\times n}$, $A + \mathbf{O} = A$.

_____ 6.If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then we have that for each A ε $\mathbf{K}^{m\times n}$, there exist a unique matrix called $-A$ such that $A + (-A) = \mathbf{O}$.

_____ 7. If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then we have $O + A = A$.

_____ 8. If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then $(-B) + (A + B) = (A + B) + (-B) = A$

_____ 9. If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then for all A, B, C $\in\mathbf{K}^{m\times n}$, we have
$-(A + B) = (-A) + (-B)$.

_____ 10. If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then for all A, B, C $\in\mathbf{K}^{m\times n}$, we have that the additive inverse of a sum is the sum of the additive inverses.

_____ 11. If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then for all A, B, C $\in\mathbf{K}^{m\times n}$, we have that if
$A + B = A + C$, then $B = C$

_____ 12. If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then $A + B = A + C$, then $B = C$ is called the cancellation law for addition.

_____ 13. Many of the properties of matrices are identies and can be proved using the STATEMENT/REASON format.

_____ 14. Some of the properties of matrices do not rely on looking at the components of the matrices.

_____ 15. Some of the properties of matrices do not rely directly on the properties of the underlying field.

_____ 16. Some of the properties of matrices are not identities.

_____ 17. Some of the properties of matrices are conditional equalities.

_____ 18. Some of the properties of matrices can be proved by starting with one side and using the substitution axiom of equality to achieve the other side.

_____ 19. Sometimes a better proof of a matrix property can be obtained by simply adding the same element to both sides of a given equation.

_____ 20. Sometimes a proofs of a matrix properties can be obtained by using properties of equality.

_____ 21. A group is an abstract algebraic structure.

_____ 22. An Abelian group is an abstract algebraic structure.

_____ 23. A commutative group is an abstract algebraic structure.

______ 24. The set of all matrices over a field **K** is a group.
______ 25. The set of all matrices over a field **K** is an abelian group.
______ 26. The set of all matrices over a field **K** is a commutative group.
______ 27. The set of all matrices over a field **K** is a group.
______ 28. If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then for all scalars α, ß $\in$ **K** and A $\in \mathbf{K}^{m\times n}$, we have α (ß A) = (α ß) A
______ 29. If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then for all scalars α, ß $\in$ **K** and A,B $\in \mathbf{K}^{m\times n}$, we have (α + ß) A = α A + ß A
______ 30. If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then for all scalars $\alpha \in$ **K** and A,B $\in \mathbf{K}^{m\times n}$, we have α (A + B) = α A + α B
______ 31. If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then for all A $\in \mathbf{K}^{m\times n}$, we have 1 A = A .
______ 32. If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then for all A $\in \mathbf{K}^{m\times n}$, we have 0A = O.
______ 33. If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then for all all scalars $\alpha \in$ **K** we have α O = O.
______ 34. If $\mathbf{K}^{m\times n}$ is the set of all matrices over **K**, then for all scalars $\alpha \in$ **K** and A $\in \mathbf{K}^{m\times n}$, we have that if α A = O, then either α = 0 or A = O.
______ 35. The set of all matrices over **K** is a vector space.

A SERIES OF CLASS NOTES TO INTRODUCE LINEAR AND NONLINEAR PROBLEMS
TO ENGINEERS, SCIENTISTS, AND APPLIED MATHEMATICIANS

LINEAR CLASS NOTES:
A COLLECTION OF HANDOUTS FOR
REVIEW AND PREVIEW
OF LINEAR THEORY
INCLUDING FUNDAMENTALS OF
LINEAR ALGEBRA

CHAPTER 3

Vector Spaces and Subspaces

Abstract linear algebra begins with the definition of a **vector space** (or **linear space**) as an abstract algebraic structure. We may view the eight properties in the definition as the fundamental axioms for **vector space theory**. The definition requires knowledge of another abstract algebraic structure, a **field** where we can always add (and subtract) as well as multiply (and divide except for zero), but nothing essential is lost if you always think of the field (of **scalars**) as being the real or complex numbers (Halmos 1958,p.1).

DEFINITION #1. A nonempty set of objects (**vectors**), V, together with an **algebraic field** (of **scalars**) **K**, and two algebraic operations (**vector addition** and **scalar multiplication**) which satisfy the algebraic properties listed below (**Laws of Vector Algebra**) comprise a **vector space**. (Following standard convention, although technically incorrect, we will usually refer to the **set of vectors** V as the vector space). The set of scalars **K** are usually either the real numbers **R** or the complex numbers **C** in which case we refer to V as a **real or complex vector space**. Let $\vec{x}$, $\vec{y}$, $\vec{z}$ ε V be any vectors and α,ß ε **K** be any scalars. Then the following must hold:

VS1)	$\vec{x} + (\vec{y} + \vec{z}) = (\vec{x} + \vec{y}) + \vec{z}$	Associativity of vector addition
VS2)	$\vec{x} + \vec{y} = \vec{y} + \vec{x}$	Commutativity of vector addition
VS3)	There exists a vector $\vec{0}$, such that for every $\vec{x} \in V$, $\vec{x} + \vec{0} = \vec{x}$	Existence of a right additive identity vector for vector addition
VS4)	For each $\vec{x} \in V$, there exist a vector, denoted by $-\vec{x}$, such that $\vec{x} + (-\vec{x}) = \vec{0}$.	Existence of a right additive inverse vector for each vector in V
VS5)	$\alpha(ß\vec{x}) = (\alpha ß)\vec{x}$	An associativity property for scalar multiplication
VS6)	$(\alpha+ß)\vec{x} = (\alpha\vec{x}) + (ß\vec{x})$	A distributive property for scalar multiplication and vector addition
VS7)	$\alpha(\vec{x} + \vec{y}) = (\alpha\vec{x}) + (\alpha\vec{y})$	Another distributive property for scalar multiplication and vector addition
VS8)	$1\vec{x} = \vec{x}$	A scaling property for scalar multiplication

These eight properties are an essential part of the definition of a vector space. In abstract algebra terminology, the first four properties establish a vector space with vector addition as an **Abelian (or commutative) group** (another abstract algebraic structure). The last four properties give rules on how vector addition and scalar multiplication must behave together.

Although technically not correct, we often refer to the set V of vectors as the vector space. Thus the **R**, $\mathbf{R}^2$,$\mathbf{R}^3$, $\mathbf{R}^n$, and $\mathbf{R}^{m\times n}$ are referred to as real vector spaces and **C**, $\mathbf{C}^n$, and $\mathbf{C}^{m\times n}$ as complex vector spaces. Also **Q**, $\mathbf{Q}^n$ and $\mathbf{Q}^{m\times n}$, and indeed, **K**, $\mathbf{K}^n$, and $\mathbf{K}^{m\times n}$ for any field are referred to as vector spaces. However, the definition of a vector space requires that its structure be given. Thus to rigorously represent a vector space, we use a 5-tuple consisting of 1) the set of vectors, 2) the set of scalars, 3) vector addition, 4) scalar multiplication, and 5) the zero vector (which is the only vector required to be in all vector spaces). Thus V = (V, **K**, +, , $\vec{0}$). (Since we use juxtaposition to indicate scalar multiplication, we have no symbol for multiplication of a vector by a scalar and hence just leave a space.)

We consider some properties that hold in all vector spaces (i.e., they are vector space properties like the original eight, but unlike the original eight, these follow from the original eight using mathematical logic). Specific examples of vector spaces are given in the next handout. Once they have been shown to be vector spaces, it is not logically necessary to show directly that the properties in Theorems #1, #2 and #3 and Corollary #4 hold in these spaces. The properties hold since we can show that they are vector space properties (i.e., they hold in an abstract vector space). However, you may wish to check out these properties in specific vector spaces (i.e., provide a direct proof) to improve your understanding of the concepts. We start with an easy property for which we provide proof.

<u>THEOREM #1.</u> Let V = (V,**K**, +, , $\vec{0}$) be a vector space. Then for all $\vec{x} \in$V, $\vec{0} + \vec{x} = \vec{x}$.

Proof. Let $\vec{x}$ be any vector in V and let $\vec{0}$ be the right additive identity for V (or V). Then

<u>STATEMENT</u>	<u>REASON</u>
$\vec{0} + \vec{x} = \vec{x} + \vec{0}$	VS2. Vector addition is commutative
$= \vec{x}$	VS3 $\vec{0}$ is the right additive identity element for V (or V).

Hence for all $\vec{x} \in$V, $\vec{0} + \vec{x} = \vec{x}$ (i.e., $\vec{0}$ is a **left additive identity element** as well as a right additive identity element).

Q.E.D.

As in the proofs of any identity, the replacement of $\vec{0} + \vec{x}$ by $\vec{x}$ is effected by the property of equality that says that in any equality, a quantity may be replaced by an equal quantity. (Note that the second equality is really $\vec{0} + \vec{x} = \vec{x}$ as the LHS is assumed.) This proof is in some sense identical to the proof that for all x$\in$**K**, 0 + x = 0 for fields. This is because the property is really a **group theory property** and vectors with vector addition (as well as scalars with scalar addition) form <u>groups</u>. We now list additional properties of a <u>vector space</u> that follow since the vectors in a vector space along with vector addition form a <u>group</u>. Since $\vec{0}$ is both a <u>left and right additive identity element</u>, we may now say that it is **an additive identity element**.

<u>THEOREM #2.</u> Let V be a vector space over the field **K**. (i.e., Let V = (V, **K**, +, , $\vec{0}$) be a vector space.) Then

1. The zero vector is unique (i.e., there is only one additive identity element in V).
2. Every vector has a unique additive inverse element.
3. $\vec{0}$ is its own additive inverse element (i.e., , $\vec{0} + \vec{0} = \vec{0}$).
4. The additive inverse of an additive inverse element is the element itself. (i.e., if $-\vec{x}$ is the additive inverse of $\vec{x}$, then $-(-\vec{x}) = \vec{x}$).

5. $-(\vec{x} = \vec{y}) = (-\vec{x}) + (-\vec{y})$. (i.e., the additive inverse of a sum is the sum of their additive inverses.)
6. Sums of vectors can be written in any order you wish.
7. If $\vec{x} + \vec{y} = \vec{x} + \vec{z}$, then $\vec{y} = \vec{z}$. (Cancellation Law for Addition)

We now give a theorem for vector spaces analogous to the one for fields that says that if the product of two numbers is zero, one of the numbers must be zero.

<u>THEOREM #3</u>. Let V be a vector space over the field **K**. (i.e., Let $V = (V, \mathbf{K}, +, , \vec{0})$ be a vector space. The scalars **K** may be thought of as either **R** or **C** so that we have a real or complex vector space, but V may not be thought of as $\mathbf{R}^n$ or $\mathbf{C}^n$.)

1. $\forall\ \vec{x} \in V$, $\quad 0\,\vec{x} = \vec{0}$.
2. $\forall\ \alpha \in \mathbf{K}$, $\quad \alpha\vec{0} = \vec{0}$
3. $\forall \alpha \in \mathbf{K}$, $\ \forall\ \vec{x} \in V$, $\quad \alpha\vec{x} = \vec{0}$ implies either $\alpha = 0$ or $\vec{x} = \vec{0}$.

<u>COROLLARY#4 (Zero Product)</u>. Let V be a vector space over the field **K**. (**K** may be thought of is either **R** or **C**.) Let $\alpha \in \mathbf{K}$, $\vec{x} \in V$. Then $\alpha\vec{x} = \vec{0}$ if and only if $\alpha = 0$ or $\vec{x} = \vec{0}$.

In the abstract algebraic definition of a vector space, vectors do not have a magnitude (or length). Later, we will discuss **normed linear spaces** where this structure is added to a vector space. However, the concept of **direction**, in the sense indicated in the following theorem, is in all vector spaces.

<u>DEFINITION #2</u>. Let V be a vector space over a field **K**. Two nonzero vectors $\vec{x}, \vec{y} \in V$ are said to be **parallel** if there is a scalar $\alpha \in \mathbf{K}$ such that $\alpha\vec{x} = \vec{y}$. (The zero vector has no direction.) If V is a real vector space so that **K**=**R**, then two (non-zero) parallel vectors are in the **same direction** if $\alpha>0$, but in **opposite directions** if $\alpha<0$.

EXERCISES on The Definition of an Abstract Vector Space

<u>EXERCISE #1</u>. True or False.

For all questions, assume V is a vector space over a field **K**, $\vec{x}, \vec{y}, \vec{z} \in V$ are any vectors and $\alpha, ß \in \mathbf{K}$ are any scalars.

_____ 1. The following property is an axiom in the definition of a vector space:
VS) $\quad \vec{x} + (\vec{y} + \vec{z}) = (\vec{x} + \vec{y}) + \vec{z}$

_____ 2. The following property is an axiom in the definition of a vector space:
VS) $\quad \vec{x} + \vec{y} = \vec{y} + \vec{x}$

_____ 3. The following property is an axiom in the definition of a vector space:

VS)There exists a vector $\vec{0}$ such that for every $\vec{x} \in V$, $\vec{x} + \vec{0} = \vec{x}$

_____ 4. The following property is an axiom in the definition of a vector space:

VS) For each $\vec{x} \in V$, there exist a vector, denoted by $-\vec{x}$, such that $\vec{x} + (-\vec{x}) = \vec{0}$.

_____ 5. The following property is an axiom in the definition of a vector space:

VS)) $\alpha(ß\vec{x}) = (\alpha ß)\vec{x}$

_____ 6. The following property is an axiom in the definition of a vector space:

VS) $(\alpha+ß)\vec{x} = (\alpha\vec{x}) + (ß\vec{x})$

_____ 7. The following property is an axiom in the definition of a vector space:

VS) $\alpha(\vec{x} + \vec{y}) = (\alpha\vec{x}) + (\alpha\vec{y})$

_____ 8. The following property is an axiom in the definition of a vector space:

VS)) $1\vec{x} = \vec{x}$

_____ 9. In V, vector addition is associative.

_____ 10. In V, vector addition is commutative.

_____ 11.In V, for vector addition, there exist a right additive identity such that for every $\vec{x} \in V$, $\vec{x} + \vec{0} = \vec{x}$.

_____ 12. For vector addition, every vector in V has a right additive inverse, denoted by $-\vec{x}$, such that $\vec{x} + (-\vec{x}) = \vec{0}$.

_____ 13. For all $\vec{x} \in V$, we have $\vec{0} + \vec{x} = \vec{x}$.

_____ 14. In V, the zero vector is unique.

_____ 15. There is only one additive identity element in V

_____ 16. In V, every vector has a unique additive inverse element.

_____ 17. In V, $\vec{0}$ is its own additive inverse element

_____ 18. The additive inverse of an additive inverse element is the element itself

_____ 19. if $-\vec{x}$ is the additive inverse of $\vec{x}$ in V, then $-(-\vec{x}) = \vec{x}$.

_____ 20. $\forall \alpha \in \mathbf{K}$, $\forall\ \vec{x} \in V$, $\alpha\vec{x} = \vec{0}$ implies either $\alpha = 0$ or $\vec{x} = \vec{0}$.

Halmos, P. R.1958, *Finite Dimensional Vector Spaces* (Second Edition) Van Nostrand Reinhold Company, New York.

If we define a specific set of vectors, a set of scalars and two operations that satisfy the eight properties in the definition of a vector space (i.e., the Laws or Axioms of Vector Algebra) we obtain an example of a vector space. Note that this requires that the eight properties given in the definition of an (abstract) vector space be verified for the (concrete) example. We give examples of vector spaces, but the verification that they are indeed vector spaces is left to the exercises. We also give two ways of building a more complicated vector space from a given vector space.

EXAMPLE #1. The set of **sequences** of a fixed length in a field **K**. Since we are not interested in matrix multiplication at this time, we may think of them not only as sequences, but as **column vectors** or as **row vectors**. When we consider linear operators, we wish them to be column vectors and so we use this interpretation. For example, if **K** = **R**, we let

i) $$V = \left\{ \vec{x} = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} : x_i \in R, i = 1,2,3,..,n \right\} = \mathbf{R}^{n\text{x}1} \cong \mathbf{R}^n$$

ii) The scalars are the **real numbers R**.

iii) **Vector addition** is defined by

$$\begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n + y_n \end{bmatrix}.$$ (i.e. Add componentwise as in matrix addition)

iv) **Scalar multiplication** is defined by

$$\alpha \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = \begin{bmatrix} \alpha x_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha x_n \end{bmatrix}.$$ (i.e. Multiply each component in x by α as in multiplication of a matrix by a scalar.)

Again, the space of **row vectors**: $V = \{(x_1,...,x_n)\colon x_i \in \mathbf{R},\ i = 1,...,n\} = \mathbf{R}^{1\text{x}n} \cong \mathbf{R}^n$ is technically different from the space of column vectors. However, as far as being a vector space is concerned, the distinction is merely technical and not substantive. Unless we need to make use of the technical difference we will denote both spaces by $\mathbf{R}^n$. We refer to $\mathbf{R}^n$ as the space of n-tuples of real numbers whether they are written as row vectors or column vectors. If we wish to think of row vectors as $n \times 1$ matrices, column vectors as $1 \times n$ matrices, and consider matrix multiplication, then we must distinguish between row and column vectors. If we wish the scalars to be the complex numbers, we must consider the space $\mathbf{C}^n$ of n-tuples of complex numbers.

EXAMPLE #2. Consider the set $\mathbf{R}^{m\times n}$ of all **matrices** of real numbers with a fixed size, $m \times n$. We define vector addition as matrix addition. (To find the sum of two matrices, add them componentwise.) Scalar multiplication of a vector (i.e. a matrix) by a scalar is defined in the usual way (multiply each component in the matrix by the scalar). Note that matrix multiplication is not required for the set of matrices of a given size to have the vector space structure. The Laws of Vector Algebra can be shown to hold and hence $\mathbf{R}^{m\times n}$ is a vector space. We may also consider $\mathbf{C}^{m\times n}$, the set of all matrices of complex numbers.

EXAMPLE #3. **Function Spaces**. Let $\mathscr{F}(D,\mathbf{R}) = \{f: D \to \mathbf{R}\}$ where $D \subseteq \mathbf{R}$; that is, let $\mathscr{F}(D)$ be the collection of all real valued functions of a real variable which have a common domain D in $\mathbf{R}$. Often, $D = I = (a,b)$. The scalers will be $\mathbf{R}$. Vector addition is defined as function addition. Recall that a function f is defined by knowing (the rule that determines) the values $y = f(x)$ for each x is the domain of f. Given two functions f and g whose domains are both D, we can define a new function $h = f + g$ (h is called the sum of f and g) by the rule $h(x) = f(x) + g(x)$ for all $x \in D$. Similarly we define the function αf by the rule $(\alpha f)(x) = \alpha\ f(x)$ for all $x \in D$. This defines scalar multiplication of a “vector” (i.e. a function) by a scalar. The Laws of Vector Algebra can be shown to hold and hence $\mathscr{F}(D)$ is a vector space. We may also define $\mathscr{F}(D)$ to be $\{f: D \to \mathbf{C}\}$ where $D \subseteq \mathbf{C}$. That is, we may also let $\mathscr{F}(D)$ describe the set of all complex valued functions of a complex variable that have a common domain D in $\mathbf{C}$. Function spaces are very important when you wish to solve differential equations.

EXAMPLE #4. Suppose $V_{\mathbf{R}}$ is a real vector space. We construct a complex vector space $V_{\mathbf{C}}$ as follows: As a set, let $V_{\mathbf{C}} = V_{\mathbf{R}} \times V_{\mathbf{R}} = \{(\vec{x}, \vec{y}): \vec{x}, \vec{y} \in V_{\mathbf{R}}\}$. However, we will use the Eulerian notation, $\vec{z} = \vec{x} + i\vec{y}$ for elements in $V_{\mathbf{C}}$. We define vector addition and scalar multiplication componentwise in the obvious way:

If $\vec{z}_1 = \vec{x}_1 + i\vec{y}_1$, $\vec{z}_2 = \vec{x}_2 + i\vec{y}_2 \in V_{\mathbf{C}}$, then $\vec{z}_1 + \vec{z}_2 = (\vec{x}_1 + \vec{x}_2) + i(\vec{y}_1 + \vec{y}_2)$.

If $\gamma=\alpha+i\beta\in\mathbf{C}$ and $\vec{z} = \vec{x} + i\vec{y} \in V_{\mathbf{C}}$, then $\gamma\ \vec{z} = (\alpha\ \vec{x} - \beta\ \vec{y}) + i(\beta\ \vec{x} + \alpha\vec{y})$.

It is straight forward to show that with these definitions, all eight of the Laws of Vector Algebra are satisfied so that $V_{\mathbf{C}}$ is a complex vector space. We see that $V_{\mathbf{R}}$ can be embedded in $V_{\mathbf{C}}$ and hence can be considered as a subset of $V_{\mathbf{C}}$. It is not a subspace (see the next handout) since they use a different set of scalars. However, if scalars are restricted to $\mathbf{R}$ and vectors to the form $\vec{z} = \vec{x} + i\vec{0}$, then the vector space structure of $V_{\mathbf{R}}$ is also embedded in $V_{\mathbf{C}}$.

It is important to note that this process can be done for any vector space $V_{\mathbf{R}}$. If we start with $\mathbf{R}$, we get $\mathbf{C}$. If we start with $\mathbf{R}^n$, we get $\mathbf{C}^n$. If we start with $\mathbf{R}^{m\times n}$, we get $\mathbf{C}^{m\times n}$. If we start with real functions of a real variable, we get complex functions of a real variable. For example, the complex vector space $C^1(\mathbf{R}^2,\mathbf{C}) = C^1(\mathbf{R}^2,\mathbf{R})+iC^1(\mathbf{R}^2,\mathbf{R})$ with vectors that are complex valued functions of two real variables of the form u(x,y) +iv(x,y) with $u,v\in C^1(\mathbf{R}^2,\mathbf{R})$ will be of interest later. Since there is a one-to-one correspondence between $\mathbf{R}^2$ and $\mathbf{C}$ (in fact, they are isomorphic as real vector spaces) and have the same topology (they have the same norm and hence the same metric, see Chapter 8) we may identify $C^1(\mathbf{R}^2,\mathbf{R})$ with set of real valued functions of a complex variable $C^1(\mathbf{C},\mathbf{R})= \{u(z)= \tilde{u}(x, y) \in C^1(\mathbf{R}^2):z=x+iy\}$ and hence $C^1(\mathbf{R}^2,\mathbf{R})+iC^1(\mathbf{R}^2,\mathbf{R})$ with the complex vector space of complex valued functions of a complex variable $C^1(\mathbf{C},\mathbf{C}) = C^1(\mathbf{C},\mathbf{R})+iC^1(\mathbf{C},\mathbf{R})$ of the form w(z) = u(z) +iv(z) where $u,v\in C^1(\mathbf{R}^2,\mathbf{R})$.

<u>EXAMPLE</u> <u>#5</u>. **Time varying vectors**. Suppose V is a real vector space (which we think of as a state space). Now let $V(I) = \{x(t):I\to V\}=\mathcal{F}(I,V)$ where $I = (a,b)\subseteq\mathbf{R}$. That is, V is the set of all "vector valued" functions on the open interval I. (Thus we allow the state of our system to vary with time.) To make V(I) into a vector space, we must equip it with a set of scalars, vector addition, and scalar multiplication. The set of scalars for V(I) is the same as the scalars for V (i.e.,$\mathbf{R}$). Vector addition and scalar multiplication are simply function addition and scalar multiplication of a function. To avoid introducing to much notation, the engineering convention of using the same symbol for the function and the dependent variable will be used (i.e., instead of y=f(x), we use y=y(x)). Hence instead of $\vec{x} = \vec{f}(t)$, for a function in V(I), we use $\vec{x} = \vec{x}(t)$. The context will explain whether $\vec{x}$ is a vector in V or a function in V(I).

1) If $\vec{x}, \vec{y} \in V(I)$, then we define $\vec{x}+\vec{y}$ pointwise as , $(\vec{x}+\vec{y})(t) = \vec{x}(t) + \vec{y}(t)$.

2) If $\vec{x} \in V(I)$ and α is a scalar, then we define $(\alpha\, \vec{x})(t) \in V(I)$ pointwise as $(\alpha\, \vec{x})(t) = \alpha\, \vec{x}(t)$.

The proof that V(I) is a vector space is left to the exercises. We use the notation V(t) instead of V(I), when, for a math model, the interval of validity is unknown and hence part of the problem. Since V is a real vector space, so is V(t). V(t) can then be embedded in a complex vector space as described above. Although we rarely think of time as being a complex variable, this is often useful mathematically to solve dynamics problems since we may wish state variables to be analytic. Thus the holomorphic function spaces are of interest.

<u>EXAMPLE</u> <u>#6</u>. (Holomorphic functions) Consider $C^1(\mathbf{R}^2,\mathbf{R})+iC^1(\mathbf{R}^2,\mathbf{R})$.

EXERCISES on Examples of Vector Spaces

EXERCISE #1. True or False.

For all questions, assume V is a vector space over a field **K**, $\vec{x}$, $\vec{y}$, $\vec{z} \in V$ are any vectors and $\alpha, ß \in \mathbf{K}$ are any scalars.

______ 1. $V = \left\{ \vec{x} = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} : x_i \in R, i = 1,2,3,..,n \right\} = \mathbf{R}^n$ is a vector space.

______ 2. The scalars for the vector space $\mathbf{R}^3$ are the real numbers **R**.

______ 3. $\mathbf{R}^{mxn}$ is a vector space

______ 4.The scalars for the vector space $\mathbf{R}^{3x2}$ are the real numbers **R**.

______ 5. $\mathbf{C}^{1x2}$ is a vector space

______ 6. $\mathbf{R}^{mxn}$ is a real vector space

______ 7. The scalars for the vector space $\mathbf{R}^{3x2}$ are the real numbers **R**.

______ 8. The scalars for the vector space $\mathbf{C}^{3x2}$ are the real numbers **R**.

______ 9. The scalars for the vector space $\mathbf{R}^{10}$ are the real numbers **R**.

______ 10. $\mathbf{R}^{1x2}$ is a real vector space.

______ 11. $\mathbf{C}^{mxn}$ is a real vector space.

______ 12. The function space $\mathscr{F}(D,\mathbf{R}) = \{f: D \to \mathbf{R}\}$ where $D \subseteq \mathbf{R}$ is a vector space.

______ 13. The set of all continuous functions on the open interrval I = (a,b), C(I,**R**), is a vector space

______ 14. The set of all analytic functions on the open interrval I = (a,b), A (I,**R**), is a vector space.

______ 15. The scalars for the vector space $\mathscr{F}(D,\mathbf{R}) = \{f: D \to \mathbf{R}\}$ where $D \subseteq \mathbf{R}$ are the real numbers **R**.

______ 16. $\mathscr{F}(D,\mathbf{R}) = \{f: D \to \mathbf{R}\}$ where $D \subseteq \mathbf{R}$ is a real vector space.

______ 17. The set of all real valued continuous functions on the open interrval I = (a,b), C(I,**R**), is a real vector space.

An important concept in abstract linear algebra is that of a subspace. After we have established a number of important examples of vector spaces, we may have reason to examine subsets of these vector spaces. Some subsets are subspaces, some are not. It is important to be able to determine if a given subset is in fact a subspace.

In general in mathematics, a set with structure is called a **space**. A subset of a space that has the same structure as the space itself is called a **subspace**. Sometimes, all subsets of a space are subspaces with the induced mathematical structure. However, all subsets of a vector (or linear) space need not be subspaces. This is because not all subsets have a **closed algebraic structure**. The sum of two vectors in a subset might not be in the subset. A scalar times a vector in the subset might not be in the subset.

DEFINITION #1. Let W be a nonempty subset of a vector space V. If for any vectors $\vec{x}, \vec{y} \in$ W and scalars $\alpha, \beta \in \mathbf{K}$ (recall that normally the set of scalars **K** is either **R** or **C**), we have that $\alpha \vec{x} + \beta \vec{y} \in$ W, then W is a **subspace** of V.

THEOREM #1. A nonempty subset W of a vector space V is a subspace of V if and only if for $\vec{x}, \vec{y} \in$ V and $\alpha \in \mathbf{K}$ (i.e. α is a scalar) we have.

i) $\vec{x}, \vec{y} \in$ W implies $\vec{x} + \vec{y} \in$ W, and

ii) $\vec{x} \in$ W implies $\alpha \vec{x} \in$ W.

TEST FOR A SUBSPACE. Theorem #1 gives a good test to determine if a given subset of a vector space is a subspace since we can test the closure properties separately. Thus if $W \subset V$ where V is a vector space, to determine if W is a subspace,we check the following three points.

1) Check to be sure that W is nonempty. (We usually look for the zero vector since if there is $\vec{x} \in$W, then $0\vec{x} = \vec{0}$ must be in W. Every vector space and every subspace must contain the zero vector.)
2) Let $\vec{x}$ and $\vec{y}$ be arbitrary elements of W and check to see if $\vec{x} + \vec{y}$ is in W. (Closure of vector addition)
3) Let $\vec{x}$ be an arbitrary element in W and check to see if $\alpha \vec{x}$ is in W. (Closure of scalar multiplication).

The checks can be done in different ways. For example, by using an English description of the subspaces, by visualizing the subspaces geometrically, and by using mathematical notation (which you may have to develop). Although, in some sense any clear argument to validate the three properties constitutes a proof, mathematicians prefer a proof using mathematical notation since this does not limit one to three dimensions, avoids the potential pitfalls of vagueness inherent in English prose, and usually is more terse. Hence we will refer to the use of English prose and the visualization of subspaces geometrically as checks, and the use of mathematical notation as a proof. Use English prose and geometry to check in your mind. Learn to write algebraic proofs using mathematical notation to improve your understanding of the concept.

EXAMPLE #1. Let $\mathbf{R}^3 = \{(x,y,z): x,y,z\in\mathbf{R}\}$ and W be the subset of V which consists of vectors whose first and second components are the same and whose third component is zero. In mathematically (more concise)notation we can define W as $W = \{(\alpha,\alpha,0): \alpha \in \mathbf{R}\}$. (Since matrix multiplication is not needed, we use row vectors rather than column vector to save space.) Note that the scalars are the reals.

English Check. We first check to see if W is a subspace using English.

1) "Clearly" W is nonempty since, for example if $\alpha = 0$, the vector (0,0,0) is in W. Recall that the zero vector must always be in a subspace.
2) If we add two vectors in W (i.e. two vectors whose first two components are equal and whose third component is zero), "clearly" we will get a vector in W (i.e. a vector whose first two components are equal and whose third component is zero). Hence W is closed under vector addition.
3) If we multiply a vector in W (i.e. a vector whose first two components are equal and whose third component is zero) by a scalar, "clearly" we will get a vector in W (i.e. a vector whose first two components are equal and whose third component is zero). Hence W is closed under scalar multiplication.

Note that if the parenthetical expressions are removed, the above "check" does not explain or prove anything. Hence, after a geometrical check, we give a more mathematical proof using the definition of W as $W = \{(\alpha,\alpha,0): \alpha \in \mathbf{R}\}$.

Geometrical Check. Since we can sketch $\mathbf{R}^3$ in the plane $\mathbf{R}^2$ we can give a geometrical interpretation of W and can in fact check to see if W is a subspace geometrically. Recall that geometrically we associate the algebraic vector $\vec{x} = (x,y,z) \in \mathbf{R}^3$ with the geometric **position vector** (directed line segment) whose tail is the origin and whose head is at the point (x,y,z). Hence W is the set of position vectors whose heads fall on the line through the origin and the point (1,1,0). Although technically incorrect, we often shorten this and say that W is the line through the origin and point (1,1,0).

1) The line W, (i.e. the line through (0,0,0) and (1,1,0)) "clearly" contains the origin (i.e. the vector $\vec{0} = (0,0,0)$).
2) If we add two vectors whose heads lie on the line W (i.e. the line through (0,0,0) and (1,1,0)), "clearly" we will get a vector whose head lies on the line W (i.e. the line through (0,0,0) and (1,1,0)). (Draw a picture.) Hence W is closed under vector addition.
3) If we multiply a vector whose head lies on the line through W (i.e. the line through (0,0,0) and (1,1,0)) by a scalar, "clearly" we will get a vector whose head lies on the line through (0,0,0) and (1,1,0). (Again, draw a picture.) Hence W is closed under scalar multiplication.

THEOREM #2. The set $W = \{(\alpha,\alpha,0): \alpha \in \mathbf{R}\}$ is a subspace of the vector space $\mathbf{R}^3$.

Proof. Let $W = \{(\alpha,\alpha,0): \alpha \in \mathbf{R}\}$.

1) Letting $\alpha = 0$ we see that $(0,0,0) \in W$ so that $W \neq \emptyset$.

2) Now let $\vec{x} = (\alpha,\alpha,0)$ and $\vec{y} = (\beta,\beta,0)$ so that $\vec{x}$, $\vec{y}$ y $\in$ W. Then

STATEMENT	REASON
$\vec{x} + \vec{y} = (\alpha,\alpha,0) + (\beta,\beta,0)$	Notation
$= (\alpha+\beta,\alpha+\beta,0)$.	Definition of Vector Addition in $\mathbb{R}^3$

"Clearly" $(\alpha+\beta,\alpha+\beta,0) \in W$ since it is of the correct form. Hence W is closed under vector addition.

3) Let $\vec{x} = (\alpha,\alpha,0)$ and $a \in \mathbf{R}$ be a scalar. Then

STATEMENT	REASON
$a\,\vec{x} = a(\alpha,\alpha,0)$	Notation
$= (a\alpha,a\alpha,0)$.	Definition of Scalar Multiplication in $\mathbb{R}^3$

"Clearly" $(a\alpha,a\alpha,0) \in W$ since it is of the correct form. Hence W is closed under scalar multiplication.

Q.E.D.

Note that the proof is in some sense more convincing (i.e. more rigorous) than the English check or a geometric check. However we did have to develop some notation to do the proof.

EXAMPLE #2. Let V = $\mathscr{F}$(I) be the set of real valued functions on I=(a,b) where a<b. Now let W=C(I) be the set of continuous functions on I. Clearly W=C(I) is a subset of V =$\mathscr{F}$(I). Note that the scalars are the reals.

English Check. We first check to see if W is a subspace using English.

1) "Clearly" W is nonempty since, for example $f(x) = x^2$ is a continuous function in C(I). Also, $f(x) = 0$ is continuous as is required for W to be a subspace.
2) If we add two continuous functions, we get a continuous function so that sums of functions in W are in W. Hence W is closed under vector addition. (Recall that this is a theorem from calculus.)
3) If we multiply a continuous function by a scalar, we get a continuous function. Hence W is closed under scalar multiplication. (Recall that this is anther theorem from calculus.)

“Clearly” there is no “geometrical” check that W is a subspace. The reason is that C(I) is infinite dimensional and hence a geometrical picture is not possible. The concept of a basis for a vector space or subspace and that of dimension will be discussed later.)

The informal English check can be upgraded to a mathematical proof by a careful rewording and by citing references for the appropriate theorems from calculus.

We close by noting that a subspace is in fact a vector space in its on right using the vector addition and scalar multiplication of the larger vector space.

THEOREM #3. Suppose $W \subseteq V$ where V is a vector space over the scalars **K**. If W is a subspace of V, then all of the eight properties hold for all elements in W and all scalars in **K** (without relying on any of the elements in V that are not in W). Thus W with the vector addition and scalar multiplication of V (but without the elements in V that are not in W) is a vector space.

EXERCISES on Subspace of a Vector Space

EXERCISE #1. True or False.

For all questions, assume V is a vector space over a field **K**, $\vec{x}$, $\vec{y}$, $\vec{z} \in V$ are any vectors and $\alpha, \beta \in \mathbf{K}$ are any scalars.

_____ 1. If $W \subseteq V$ is a nonempty subset of a vector space V and for any vectors $\vec{x}$, $\vec{y} \in W$ and scalars $\alpha, \beta \in \mathbf{K}$, we have that $\alpha\vec{x} + \beta\vec{y} \in W$, then W is a subspace of V.

_____ 2. A nonempty subset W of a vector space V is a subspace of V if and only if for $\vec{x}$, $\vec{y} \in V$ and $\alpha \in \mathbf{K}$ (i.e. α is a scalar) we have i) $\vec{x}$, $\vec{y} \in W$ implies $\vec{x} + \vec{y} \in W$, and ii) $\vec{x} \in W$ implies $\alpha\vec{x} \in W$.

_____ 3. If W is a subspace of V, then W is nonempty.

_____ 4. If W is a subspace of V, then for any arbitrary vectors $\vec{x}$ and $\vec{y}$ in W, the sum $\vec{x} + \vec{y}$ is in W.

_____ 5. If W is a subspace of V, then for any arbitrary vector $\vec{x}$ in W and scalar $\alpha \in \mathbf{K}$, $\alpha\vec{x}$ is in W.

_____ 6. If W be the subset of V= $\mathbf{R}^3$ which consists of vectors whose first and second components are the same and whose third component is zero, then W is a subspace of V.

_____ 7. $W = \{(\alpha,\alpha,0) : \alpha \in \mathbf{R}\}$ is a subspace of $\mathbf{R}^3$.

_____ 8. $W = \{(\alpha,0,0) : \alpha \in \mathbf{R}\}$ is a subspace of $\mathbf{R}^3$.

_____ 9. If I = (a,b), then C(I,**R**) is a subspace of $\mathcal{F}$ (I,**R**).

_____ 10. If I = (a,b), then $\mathcal{A}$ (I,**R**) is a subspace of $\mathcal{F}$ (I,**R**).

_____ 11. If I = (a,b), then $\mathcal{A}$ (I,**R**) is a subspace of C(I,**R**).

EXERCISE #2. Let $\mathbf{R}^3 = \{(x,y,z) : x,y,z \in \mathbf{R}\}$ and W be the subset of V which consists of vectors where all three components are the same. In mathematically (more concise) notation we can define W as $W = \{(\alpha,\alpha,\alpha) : \alpha \in \mathbf{R}\}$. (Since matrix multiplication is not needed, we use row vectors rather than column vector to save space.) Note that the scalars are the reals. If possible, do an English check, a geometric check, and a proof that W is a subspace.

EXERCISE #3. Let $\mathbf{R}^3 = \{(x,y,z) : x,y,z \in \mathbf{R}\}$ and W be the subset of V which consists of vectors where second ands third components are zero. In mathematically (more concise) notation we can define W as $W = \{(\alpha,0,0) : \alpha \in \mathbf{R}\}$. (Since matrix multiplication is not needed, we use row vectors rather than column vector to save space.) Note that the scalars are the reals. If possible, do an Englishcheck, a geometric check, and a proof that W is a subspace.

EXERCISE #4. Let $\mathbf{R}^3 = \{(x,y,z): x,y,z\in\mathbf{R}\}$, $\mathbf{R}^4 = \{(x_1,x_2,x_3, x_4): x_1,x_2,x_3, x_4\in\mathbf{R}\}$. Consider the following subsets W of these vector sapaces. Determine which are subspaces. (Since matrix multiplication is not needed, we use row vectors rather than column vector to save space.) Note that the scalars are the reals. For each subset, first write an English description of the subset. If possible, do an English check, a geometric check, and a proof that W is a subspace. If W is not a subspace explain clearly why it is not. In mathematically (more concise) notation we define W as

1. $W = \{(\alpha,1,0): \alpha \in \mathbf{R}\} \subseteq\mathbf{R}^3$.
2. $W = \{(x,y,z): x^2 + y^2 + z^2 =1 \text{ and } x,y,z\in \mathbf{R}\} \subseteq\mathbf{R}^3$
3. $W = \{(\alpha,\beta,0): \alpha,\beta \in \mathbf{R}\}\subseteq\mathbf{R}^3$.
4. $W = \{(\alpha,\beta,0,0): \alpha,\beta \in \mathbf{R}\}\subseteq\mathbf{R}^4$.
5. $W = \{(\alpha,\alpha,\alpha,\alpha): \alpha \in \mathbf{R}\}\subseteq\mathbf{R}^4$
6. $W = \{(\alpha,\alpha,\beta,\beta): \alpha,\beta \in \mathbf{R}\}\subseteq\mathbf{R}^4$

A SERIES OF CLASS NOTES TO INTRODUCE LINEAR AND NONLINEAR PROBLEMS TO ENGINEERS, SCIENTISTS, AND APPLIED MATHEMATICIANS

LINEAR CLASS NOTES:
A COLLECTION OF HANDOUTS FOR
REVIEW AND PREVIEW
OF LINEAR THEORY
INCLUDING FUNDAMENTALS OF
LINEAR ALGEBRA

CHAPTER 4

Introduction to Solving

Linear Algebraic Equations

In high school you learned how to solve two linear (and possibly nonlinear) equations in two unknowns by the elementary algebraic techniques of **addition** and **substitution** to eliminate one of the variables. These techniques may have been extended to three and four variables. Also, **graphical** or **geometric techniques** may have been developed using the intersection of two lines or three planes. Later, you may have been introduced to a more formal approach to solving a system of m equations in n unknown variables.

$$\begin{array}{c} a_{11}\ x_1 + a_{12}\ x_2 + \cdots + a_{1n}\ x_n = b_1 \\ a_{21}\ x_1 + a_{22}\ x_2 + \cdots + a_{2n}\ x_n = b_2 \\ \vdots \quad\quad \vdots \quad\quad \vdots \quad\quad \vdots \\ a_{m1}\ x_1 + a_{m2}\ x_2 + \cdots + a_{mn}\ x_n = b_m \end{array} \tag{1}$$

We assume all of the **scalars** a_{ij}, x_i, and b_j are elements in a **field K**. (Recall that a **field** is an abstract algebra structure that may be defined informally as a number system where you can add, subtract, multiply , and divide. Examples of a field are **Q**, **R**, and **C**. However, **N** and **Z** are not fields. Unless otherwise noted, the scalars can be assumed to be real or complex numbers. i.e., **K** = **R** or **K** = **C**.) The formal process for solving **m linear algebraic equations in n unknowns** is called **Gauss Elimination**. We need a formal process to prove that a solution (or a parametric expression for an infinite number of solutions) can always be obtained in a finite number of steps (or a proof that no solution exists), thus avoiding the pitfall of a "circular loop" which may result from ad hoc approaches taken to avoid fractions. Computer programs using variations of this algorithm avoid laborious arithmetic and handle problems where the number of variables is large. Different programs may take advantage of particular characteristics of a category of linear algebraic equations (e.g., banded equations). Software is also available for **iterative techniques** which are not discussed here. Another technique which is theoretically interesting, but only useful computationally for very small systems is **Cramer Rule** which is discussed in Chapter 7.

DEFINITION #1. A **solution** to (1) is an n-tuple (finite sequence) $x_1, x_2, ..., x_n$ in $\mathbf{K}^n$ (**e.g.,** $\mathbf{R}^n$ or $\mathbf{C}^n$) such that all of the equations in (1) are true. It can be considered to be a row vector $[\ x_1, x_2, ..., x_n\ ]$ or as a column vector $[\ x_1, x_2, ..., x_n\ ]^T$ using the transpose notation. When we later formulate the problem given by the scalar equatiuons (1) as a matrix or "vector" equation, we will need our unknown "vector" to be a column vector, hence we use column vectors. If we use column vectors, the **solution set** for (1) is the set

$S = \{ \vec{x} = [\ x_1, x_2, ..., x_n\ ] \in \mathbf{K}^n$: when $x_1, x_2, ..., x_n$ are substituted into (1), all of the equations in (1) are satisfied$\}$.

DEFINITION #2. Two systems of linear algebraic equations are **equivalent** if their solution sets are equal (i.e., have the same elements).

As with a single algebraic equation, there are algebraic operations that can be performed on a system that yield a new equivalent system. We also refer to these as **equivalent equation operations** (EEO's). However, we will restrict the EEO's we use to three **elementary equation operations**.

DEFINITION #3. The **Elementary Equation Operations (EEO's)** are
1. Exchange two equations
2. Multiply an equation by a nonzero constant.
3. Replace an equation by itself plus a scalar multiple of another equation.

The hope of elimination using EEO's is to obtain, if possible, an equivalent system of n equations that is **one-way coupled** with new coefficients a_{ij} as follows:

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &=& b_1 \\ a_{22}x_2 + \cdots + a_{2n}x_n &=& b_2 \\ \vdots & & \vdots \\ a_{nn}x_n &=& b_n \end{array} \qquad (2)$$

where $a_{ii} \neq 0$ for i = 1, 2, ...,n. Since these equations are only one-way coupled, the last equation may then be solved for x_n and substituted back into the previous equation to find x_{n-1}. This process may be continued to find all of the x_i's that make up the unique (vector) solution. Note that this requires that $m \geq n$ so that there are at least as many equations as unknowns (some equations may be redundant) and that all of the diagonal coefficients a_{ii} are not zero. This is a very important special case. The nonzero coefficients a_{ii} are called **pivots**.

Although other operations on the system of equations can be derived, the three EEO's in Definition #3 are sufficient to find a one-way coupled equivalent system, if one exists. If sufficient care is not taken when using other operations such as replacing an equation by a linear combination of the equations, a new system which is not equivalent to the old one may result. Also, if we restrict ourselves to these three EEO's, it is easier to develop computational algorithms that can be easily programed on a computer. Note that applying one of these EEO's to a system of equations (the Original System or OS) results in a new system of equations (New System or NS). Our claim is that the systems OS and NS have the same solution set.

THEOREM #1. The new system (NS) of equations obtained by applying an elementary equation operation to a system (OS) of equations is equivalent to the original system (OS) of equations.

Proof idea. By Definition#2 we need to show that the two systems have the same solution set.

EEO # 1 Whether a given $\vec{x} = [x_1, x_2, ..., x_n]$ is a solution of a given scalar equation is "clearly" not dependent on the order in which the equations are written down. Hence whether a given $\vec{x} = [x_1, x_2, ..., x_n]$ is a solution of all equations in a system of equations is not dependent on the order in which the equations are written down

EEO # 2 If $\vec{x} = [x_1, x_2, ..., x_n]$ satisfies.

$$a_{i1}x_1 + ka_{i2}x_2 + \cdots + a_{in}x_n = b_i \tag{3}$$

then by the theorems of high school algebra, for any scalar k it also satisfies

$$ka_{i1}x_1 + ka_{i2}x_2 + \cdots + ka_{in}x_n = kb_{i.} \tag{4}$$

The converse can be shown if $k \neq 0$ since k will have a multiplicative inverse.

EEO # 3. If $\vec{x} = [x_1, ..., x_n]^T$ satisfies each of the original equations, then it satisfies

$$a_{i1}x_1 + \cdots + a_{in}x_n + k(a_{j1}x_1 + \cdots + a_{jn}x_n) = b_i + kb_j. \tag{5}$$

Conversely, if (5) is satisfied and

$$a_{j1}x_1 + \cdots + a_{jn}x_n = b_j \tag{6}$$

is satisfied then (5) - k (6) is also satisfied. But this is just (3).

QED

<u>THEOREM #2</u>. Every EEO has in inverse EEO of the same type.

Proof idea. EEO # 1. The inverse operation is to switch the equations back.
EEO # 2. The inverse operation is to divide the equation by $k(\neq 0)$; that is, to multiply the equation by 1/k.
EEO # 3. The inverse operation is to replace the equation by itself minus k (or plus $-k$) times the previously added equation (instead of plus k times the equation).

QED

EXERCISES on Systems, Solutions, and Elementary Equation Operations

<u>EXERCISE #1</u>. True or False.

_____ 1. The formal process for solving m linear algebraic equations in n unknowns is called Gauss Elimination

_____ 2. Another technique for solving n linear algebraic equations in n unknowns is Cramer Rule

_____ 3. A solution to a system of m linear algebraic equations in n unknowns is an n-tuple (finite sequence) $x_1, x_2, ..., x_n$ in $\mathbf{K}^n$ (e.g., $\mathbf{R}^n$ or $\mathbf{C}^n$) such that all of the equations are true.

Gauss elimination can be used to solve a system of m linear algebraic equations over a field **K** in n unknown variables $x_1, x_2, ..., x_n$ in a finite number of steps.

$$\begin{array}{ccccccc} a_{11}\ x_1 & + & a_{12}\ x_2 & + \cdots + & a_{1n}\ x_n & = & b_1 \\ a_{21}\ x_1 & + & a_{22}\ x_2 & + \cdots + & a_{2n}\ x_n & = & b_2 \\ \cdot & & \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot & & \cdot \\ a_{m1}\ x_1 & + & a_{m2}\ x_2 & + \cdots + & a_{mn}\ x_n & = & b_m \end{array} \tag{1}$$

Recall

DEFINITION #1. A **solution** of to (1) is an n-tuple $x_1, x_2, ..., x_n$ which may be considered as a (row) vector $[\ x_1, x_2, ..., x_n\ ]$ (or column vector) in $\mathbf{K}^n$ (usually **K** is **R** or **C**) such that all of the equations in (1) are true. The **solution set** for (1) is the set $S = \{\ [\ x_1, x_2, ..., x_n\ ] \in \mathbf{K}^n$: all of the equations in (1) are satisfied$\}$. That is, the solution set is the set of all solutions. The Σ set is the set where we look for solutions. In this case it is $\mathbf{K}^n$. Two systems of linear algebraic equations are **equivalent** if their solution set are the same (i.e., have the same elements).

Recall also the three **elementary equation operations** (EEO's) that can be used on a set of linear equations which do not change the solution set.
1. Exchange two equations
2. Multiply an equation by a nonzero constant.
3. Replace an equations by itself plus a scalar multiple of another equation.

Although other operations on the system of equations can be derived, if we restrict ourselves to these operations, it is easier to develop computational algorithms that can be easily programed on a computer.

DEFINITION #2. The **coefficient matrix** is the array of coefficients for the system (not including the right hand side, RHS).

$$\begin{bmatrix} a_{11} & a_{12} & \cdot & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{m1} & a_{m21} & \cdot & \cdot & \cdot & a_{mn} \end{bmatrix} \tag{2}$$

DEFINITION #3. The **augmented (coefficient) matrix** is the coefficient matrix augmented by the values from the right hand side (RHS).

$$\begin{array}{cccccl} x_1 & x_2 & & x_n & & \\ \left[\begin{array}{c} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{array}\right. & \begin{array}{c} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{array} & \begin{array}{c} \cdots \\ \cdots \\ \\ \cdots \end{array} & \begin{array}{c} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{array} & \left|\begin{array}{c} b_1 \\ b_2 \\ \vdots \\ b_m \end{array}\right] & \begin{array}{l} \text{Represents the first equation} \\ \text{Represents the second equation} \\ \\ \text{Represents the } m^{th} \text{ equation} \end{array} \end{array} \tag{3}$$

Given the coefficient matrix A and the right hand side (“vector”) $\vec{b}$, we denote the associated augmented matrix as $[A|\,\vec{b}\,]$. (Note the slight abuse of notation since A actually includes the brackets. This should not cause confusion.)

ELEMENTARY ROW OPERATIONS. All of the information contained in the equations of a linear algebraic system is contained in the augmented matrix. Rather than operate on the original equations using the elementary equation operations (EEO’s) listed above, we operate on the augmented matrix using **Elementary Row Operations** (ERO's).

DEFINITION #4. We define the **Elementary Row Operations (ERO's)**

1. Exchange two rows (avoid if possible since the determinant of the new coefficient matrix changes sign).
2. Multiply a row by a non zero constant (avoid if possible since the determinant of the new coefficient matrix is a multiple of the old one).
3. Replace a row by itself plus a scalar multiple of another row (the determinant of the new coefficient matrix is the same as the old one).

There is a clear one-to-one correspondence between systems of linear algebraic equations and augmented matrices including a correspondence between EEO’s and ERO’s. We say that the two structures are **isomorphic**. Using this correspondence and the theorems on EEO’s from the previous handout, we immediately have the following theorems.

THEOREM #1. Suppose a system of linear equations is represented by an augmented matrix which we call the original augmented matrix (OAM) and that the OAM is operated on by an ERO to obtain a new augmented matrix (NAM). Then the system of equations represented by the new augmented matrix (NAM) has the same solution set as the system represented by the original augmented matrix (OAM).

THEOREM #2. Every ERO has an inverse ERO of the same type.

DEFINITION #5. If A and B are $m \times n$ (coefficient or augmented) matrices over the field F, we say that B is row-equivalent to A if B can be obtained from A by finite sequence of ERO's.

THEOREM #3. Row-equivalence is an **equivalence relation** (Recall the definition of equivalence relation given in the remedial notes or look up it up in a Modern Algebra text).

THEOREM #4 If A and B are m×n augmented matrices which are row-equivalent, then the systems they represent are equivalent (i.e., have the same solution set).

Proof idea. Suppose $A = A_0 \to A_1 \to \cdots \to A_k = B$. Using induction and Theorem #1, the two systems can be shown to have the same solution set.

DEFINITION #6. If A is a m×n (coefficient or augmented) matrices over the field F, we call the first nonzero entry in each row its **leading entry**. We say that A is in **row-echelon form** if:
1. All entries below a leading entry are zero,
2. For i =2, ..., m, the leading entry in row i is to the right of the leading entry in row (i−1),
3. All rows containing only zeros are below any rows containing nonzero entries.

If, in addition,
4. The leading entry in each row is one,
5. All entries above a leading entry are zero,

then A is in **reduced-row-echelon** form (or **row-reduced-echelon** form).
If A is in row-echelon form (or reduced-row-echelon form) we refer to the leading entries as **pivots**.

For any matrix A, Gauss Elimination (GE) will always obtain a row-equivalent matrix U that is in row-echelon form. Gauss-Jordan Elimination (GJE) will yield a row-equivalent matrix R that is in reduced-row-echelon form. $[A|\vec{b}] \overset{\text{GE}}{\Rightarrow} [U|\vec{c}] \overset{\text{GJE}}{\Rightarrow} [R|\vec{d}]$

EXAMPLE #1. To illustrate Gauss elimination we consider an example:

$$2x + y + z = 1 \qquad (1)$$
$$4x + y = -2 \qquad (2)$$
$$-2x + 2y + z = 7 \qquad (3)$$

It does not illustrate the procedure completely, but is a good starting point. The **solution set** S for (1), (2) and (3) is the set of all ordered triples, [x,y,z] which satisfy all three equations. That is,

$$S = \{ \vec{x} = [x, y, z] \in \mathbf{R}^3 : \text{Equations (1), (2), and (3) are true} \}$$

The coefficient and augmented matrices are

$$
A = \begin{array}{c} \begin{array}{ccc} x & y & z \end{array} \\ \begin{bmatrix} 2 & 1 & 1 \\ 4 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix} \end{array}
\qquad
[A|\vec{b}\,] = \begin{array}{c} \begin{array}{cccc} x & y & z & \text{RHS} \end{array} \\ \left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 4 & 1 & 0 & -2 \\ -2 & 2 & 1 & 7 \end{array}\right] \end{array}
\begin{array}{l} \text{represents equation 1} \\ \text{represents equation 2} \\ \text{represents equation 3} \end{array}
$$

Note that matrix multiplication is not necessary for the process of solving a system of linear algebraic equations. However, you may be aware that we can use matrix multiplication to write the system (1), (2), and (3) as A $\vec{x} = \vec{b}$ where now it is mandatory that $\vec{x}$ be a column vector instead of a row vector. Also $\vec{b}$ = $[1, -2, 7]^T$ is a column vector. (We use the transpose notation to save space and paper.)

<u>GAUSS ELIMINATION</u>. We now use the elementary row operations (ERO's) on the example in a systematic way known as (naive) **<u>Gauss</u> (Jordan) Elimination** to solve the system. The process can be divided into three (3) steps.

<u>Step 1</u>. **Forward Elimination** (obtain zeros below the pivots).This step is also called the (forward) sweep phase.

$$
\begin{array}{r} \\ R_2 - 2R_1 \\ R_3 + R_1 \end{array}
\left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 4 & 1 & 0 & -2 \\ -2 & 2 & 1 & 7 \end{array}\right]
\Rightarrow
\begin{array}{r} \\ \\ R_3 + 3R_2 \end{array}
\left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & -1 & -2 & -4 \\ 0 & 3 & 2 & 8 \end{array}\right]
\Rightarrow
\left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & -1 & -2 & -4 \\ 0 & 0 & -4 & -4 \end{array}\right]
$$

This completes the forward elimination step. The **pivots** are the diagonal elements 2, -1, and -4. Note that in getting zeros below the pivot 2, we can do 2 ERO’s and only rewrite the matrix once. The last augmented matrix represents the system.

$2x + y + z = 1$	We could now use	$-4z = -4 \Rightarrow z = 1$
$-y - 2z = -4$	**<u>back</u> substitution**	$-y = -4 + 2z = -4 + 2(1) = -2 \Rightarrow y = 2$
$-4z = -4$.	to obtain	$2x = 1 - y - z = 1 - (2) - (1) = -2 \Rightarrow x = -1.$

The unique solution is sometimes written in the scalar form as $x = -1$, $y = 2$, and $z = 1$, but is more correctly written in the vector form as the column vector $[-1,2,1]^T$. Instead of back substituting with the equations, we can use the following two additional steps using the augmented matrix. When these steps are used, we refer to the process as **Gauss-Jordan elimination**.

<u>Step 2</u>. **Gauss-Jordan Normalization** (Obtain ones along the diagonal).

$$\begin{array}{r} 1/2R_1 \\ -R_2 \\ -1/4R_3 \end{array}\left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & -1 & -2 & -4 \\ 0 & 0 & -4 & -4 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 1/2 & 1/2 & 1/2 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

We could now use back substitution but instead we can proceed with the augmented matrix.

Step 3. **Gauss-Jordan Completion Phase** (Obtain zeros above the pivots)
Variation #1 Right to Left.

$$\begin{array}{r} nR_1 - 1/2nR_3 \\ R_2 - 2R_3 \\ \end{array}\left[\begin{array}{ccc|c} 1 & 1/2 & 1/2 & 1/2 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 1 \end{array}\right] \Rightarrow \begin{array}{r} nR_1 - 1/2nR_2 \\ \\ \end{array}\left[\begin{array}{ccc|c} 1 & 1/2 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right] \Rightarrow \begin{array}{l} x = -1 \\ y = 2 \\ z = 1 \end{array}$$

Variation #2 (Left to Right)

$$\begin{array}{r} nR_1 - 1/2\,R_2 \\ \\ \end{array}\left[\begin{array}{ccc|c} 1 & 1/2 & 1/2 & 1/2 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 1 \end{array}\right] \Rightarrow \begin{array}{r} nR_1 + 1/2nR_3 \\ R_2 - 2R_3 \\ \end{array}\left[\begin{array}{ccc|c} 1 & 0 & -1/2 & 3/2 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 1 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right] \Rightarrow \begin{array}{l} x = -1 \\ y = 2 \\ z = 1 \end{array}$$

Note that both variations as well as back substitution result in the same solution, $\vec{x} = [-1, 2, 1]^T$. It should also be reasonably clear that this algorithm can be programmed on a computer. It should also be clear that the above procedure will give a unique solution for n equations in n unknowns in a finite number of steps provided **<u>zeros never appear on the diagonal</u>**. The case when zeros appear on the diagonal and we still have a unique solution is illustrated below. The more general case of m equations in n unknowns where three possibilities exist is discussed later.

<u>EXAMPLE #2</u>. To illustrate the case of zeros on the diagonal that are eliminated by row exchanges consider:

$$-y - 2z = -4 \qquad (4)$$
$$-4z = -4 \qquad (5)$$
$$2x + y + z = 1 \qquad (6)$$

The augmented matrix is

$$[A|\vec{b}] = \begin{array}{c} \begin{array}{cccc} x & y & z & \text{RHS} \end{array} \\ \left[\begin{array}{ccc|c} 0 & -1 & -2 & -4 \\ 0 & 0 & -4 & -4 \\ 2 & 1 & 1 & 1 \end{array}\right] \end{array} \begin{array}{l} \\ \text{representsnequationn1} \\ \text{representsnequationn2} \\ \text{representsnequationn3} \end{array}$$

Note that there is a zero in the first row first column so that Gauss Elimination temporarily breaks down. However, note that the augmented matrix is row equivalent to those in the previous problem since it is just the matrix at the end of the forward step of that problem with the rows in a different order To establish a standard convention for fixing the breakdown, we go down the first column until we find a nonzero number. We then switch the first row with the first row with a nonzero entry in the first column. (If all of the entries in the first column are zero, then x can be anything and is not really involved in the problem.) Switching Rows one and three we obtain

$$\begin{array}{c}\to\\ \\ \to\end{array}\left[\begin{array}{ccc|c} 0 & -1 & -2 & -4 \\ 0 & 0 & -4 & -4 \\ 2 & 1 & 1 & 1 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & 0 & -4 & -4 \\ 0 & -1 & -2 & -4 \end{array}\right]$$

The 2 in the first row, first column is now our first pivot. We go to the second row, second column. Unfortunately, it is also a zero. But the third row, second column is not so we switch rows.

$$\begin{array}{c} \\ \to\\ \to\end{array}\left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & 0 & -4 & -4 \\ 0 & -1 & -2 & -4 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & -1 & -2 & -4 \\ 0 & 0 & -4 & -4 \end{array}\right]$$

We now have the same augmented matrix as given at the end of the forward step for the previous problem. Hence the solution is $x = -1$, $y = 2$, and $z = 1$. This can be written as the column vector $\mathbf{x} = [-1, 2, 1]^T$.

EXERCISES on Introduction to Gauss Elimination

<u>EXERCISE #1</u>. True or False.

_____ 1. An elementary equation operation (EEO) that can be used on a set of linear algebraic equations which does not change the solution set is “Exchange two equations”.

_____ 2. An elementary equation operation (EEO) that can be used on a set of linear algebraic equations which does not change the solution set is “Multiply an equation by a nonzero constant”.

_____ 3. An elementary equation operation (EEO) that can be used on a set of linear algebraic equations which does not change the solution set is “Replace an equations by itself plus a scalar multiple of another equation”.

_____ 4. An elementary row operation (ERO) of type one that can be used on a matrix is “Exchange two rows”.

_____ 5. An elementary row operation (ERO) of type two that can be used on a matrix is “Multiply a row by a nonzero constant”

_____ 6. An elementary row operation (ERO) of type three that can be used on a matrix is “Replace a row by itself plus a scalar multiple of another row”.

_____ 7. There is a clear one-to-one correspondence between systems of linear algebraic equations and augmented matrices including a correspondence between EEO’s and ERO’s so that we say that the two structures are isomorphic.

_____ 8. Every ERO has an inverse ERO of the same type.

_____ 9. If A and B are m×n (coefficient or augmented) matrices over the field **K**, we say that B is row-equivalent to A if B can be obtained from A by finite sequence of ERO's.

_____ 10. Row-equivalence is an equivalence relation

_____ 11. If A and B are m×n augmented matrices which are row-equivalent, then the systems they represent are equivalent (i.e., have the same solution set)

_____ 12. If A is a m×n (coefficient or augmented) matrix over the field **K**, we call the first nonzero entry in each row its leading entry.

EXERCISE #2. Solve
$$\begin{aligned} 4x + y \phantom{{}+z} &= -2 \\ -2x + 2y + z &= 7 \\ 2x + y + z &= 1 \end{aligned}$$

EXERCISE #3. Solve
$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 1 \\ x_1 + 2x_2 + x_3 &= 0 \\ x_3 + x_4 &= 1 \\ x_2 + 2x_3 + x_4 &= 1 \end{aligned}$$

EXERCISE #4. Solve
$$\begin{aligned} 4x + y \phantom{{}+z} &= -2 \\ -2x + y + z &= 7 \\ 2x + y + z &= 1 \end{aligned}$$

EXERCISE #5. Solve
$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 1 \\ x_1 + 2x_2 + x_3 &= 0 \\ x_3 + x_4 &= 1 \\ x_2 + 2x_3 + 2x_4 &= 3 \end{aligned}$$

EXERCISE #6. Solve
$$\begin{aligned} 4x + y \phantom{{}+z} &= -2 \\ -2x + 2y + z &= 5 \\ 2x + y + z &= 1 \end{aligned}$$

EXERCISE #7. Solve
$$\begin{aligned} 4x + y \phantom{{}+z} &= -2 \\ -2x + 2y + z &= 1 \\ 2x + y + z &= 1 \end{aligned}$$

By using the definition of **matrix equality** we can think of the system of **scalar equations**

$$\begin{array}{l} a_{11}\ x_1 + a_{12}\ x_2 + \cdots + a_{1n}\ x_n = b_1 \\ a_{21}\ x_1 + a_{22}\ x_2 + \cdots + a_{2n}\ x_n = b_2 \\ \quad\vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \qquad \vdots \\ a_{m1}\ x_1 + a_{m2}\ x_2 + \cdots + a_{mn}\ x_n = b_m \end{array} \tag{1}$$

as one **"vector" equation** where "vector" means an n-tuple or column vector. By using matrix multiplication (1) can be written as

$$\underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{b}} \tag{2}$$

where

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = [x_i] \in \mathbf{K}^n, \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{bmatrix} = [b_i] \in \mathbf{K}^m, \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdot\ \cdot\ \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot\ \cdot\ \cdot & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1} & a_{m21} & \cdot\ \cdot\ \cdot & a_{mn} \end{bmatrix} = [a_{ij}] \in \mathbf{K}^{mxn}.$$

If we think of $\mathbf{K}^n$ and $\mathbf{K}^m$ as vector spaces, we can define

$$T(\vec{x}) = \underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}} \tag{3}$$

so that T is a mapping from $\mathbf{K}^n$ to $\mathbf{K}^m$. We write T: $\mathbf{K}^n \to \mathbf{K}^m$. A mapping from a vector space to another vector space is called an **operator**. We may now view the system of scalar equations as the operator equation:

$$T(\vec{x}) = \vec{b} \tag{4}$$

A column vector $\vec{x} \in \mathbf{K}^n$ is a solution to (4) if and only if it is a solution to (1) or (2). Solutions to (4) are just those vectors $\vec{x}$ in $\mathbf{K}^n$ that get mapped to $\vec{b}$ in $\mathbf{K}^m$ by the operator T. The equation

$$T(\vec{x}) = \vec{0} \tag{5}$$

is called **homogeneous** and is the **complementary equation** to (4). Note that it always has

$\vec{x} = \vec{0} = [0, 0, \dots, 0]^T \in \mathbf{K}^n$ as a solution. $\vec{x} = \vec{0}$ is called the **trivial solution** to (5) since it is always a solution. The question is "Are there other solutions?" and if so "How many?". We now have a connection between solving linear algebraic equations, matrix algebra, and abstract linear algebra. Also we now think of solving linear algebraic equations as an example of a **mapping problem** where the operator $T:\mathbf{R}^n \to \mathbf{R}^m$ is defined by (3) above. We wish to find all solutions to the **mapping problem** (4). To introduce this topic, we define what we mean by a **linear operator** from one vector space to another.

DEFINITION #1. An operator $T:V \to W$ is said to be **linear** if $\forall\ \vec{x}, \vec{y} \in V$ and scalars α,β, it is true that

$$T(\alpha\,\vec{x} + \beta\,\vec{y}) = \alpha\,T(\vec{x}) + \beta\,T(\vec{y}). \qquad (6)$$

$T: \mathbf{K}^n \to \mathbf{K}^m$ defined by $T(\vec{x}) = \underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}}$ is one example of a linear operator. We will give others later. To connect the mapping problem to matrix algebra we reveal that if m = n, then (1) (2) and (4) have a unique solution if and only if the matrix A is invertible.

THEOREM #1. Suppose m = n so that (1) has the same number of equations as unknowns. Then (1), (2), and (4) have a unique solution if and only if the matrix A is invertible.

EXERCISES on Connection with Matrix Algebra and Abstract Linear Algebra

EXERCISE #1. True or False.

_____ 1. Since $\mathbf{K}^n$ and $\mathbf{K}^m$ are vector spaces, we can define $T(\vec{x}) = \underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}}$ so that T is a mapping from $\mathbf{K}^n$ to $\mathbf{K}^m$.

_____ 2. A mapping from a vector space to another vector space is called an operator.

_____ 3. Solutions to $\underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{b}}$ are just those vectors $\vec{x}$ in $\mathbf{K}^n$ that get mapped to $\vec{b}$ in $\mathbf{K}^m$ by the operator $T(\vec{x}) = \underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}}$.

_____ 4. The equation $T(\vec{x}) = \vec{0}$ is called homogeneous and is the complementary equation to $\underset{m\times n}{T}\left(\underset{n\times 1}{\vec{x}}\right) = \underset{m\times 1}{\vec{b}}$ where $T(\vec{x}) = \underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}}$.

_____ 5. $\vec{x} = \vec{0}$ is called the trivial solution to the complementary equation to $\underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{b}}$.

_____ 6. An operator $T:V \to W$ is said to be linear if $\forall\ \vec{x}, \vec{y} \in V$ and scalars α,β, it is true that $T(\alpha\,\vec{x} + \beta\,\vec{y}) = \alpha\,T(\vec{x}) + \beta\,T(\vec{y})$.

_____ 7. The operator T: $\mathbf{K}^n$ to $\mathbf{K}^m$ defined by $T(\vec{x}) = \underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}}$ is a linear operator.

_____ 8. $\underset{n\times n}{A}\ \underset{n\times 1}{\vec{x}} = \underset{n\times 1}{\vec{b}}$ has a unique solution if and only if the matrix A is invertible.

We consider the three possible outcomes of applying Gauss elimination to the system of scalar algebraic equations:

$$\begin{array}{l} a_{11}\ x_1 + a_{12}\ x_2 + \cdots + a_{1n}\ x_n = b_1 \\ a_{21}\ x_1 + a_{22}\ x_2 + \cdots + a_{2n}\ x_n = b_2 \\ \quad\vdots \qquad\quad \vdots \qquad\qquad \vdots \qquad \vdots \\ a_{m1}\ x_1 + a_{m2}\ x_2 + \cdots + a_{mn}\ x_n = b_m \end{array} \tag{1}$$

THEOREM. There exists three possibilities for the system (1):

1. There is no solution.
2. There exists exactly one solution.
3. There exists an infinite number of solutions

We have illustrated the case where there exists exactly one solution. We now illustrate the cases where there exists no solution and an infinite number of solutions.

EXAMPLE. Solve

$$\begin{array}{lr} \text{E1} & x + \ \ y + \ z = b_1 \\ \text{E2} & 2x + 2y + 3z = b_2 \\ \text{E3} & 4x + 4y + 8z = b_3 \end{array} \tag{2}$$

if $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ has the values 1) $\vec{b} = \vec{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and 2) $\vec{b} = \vec{b}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

Solution: Forward Step. Reduce the augmented matrix

$$\begin{array}{c} \\ R_2 - 2R_1 \\ R_3 - 4R_1 \end{array}\left[\begin{array}{ccc|c} 1 & 1 & 1 & b_1 \\ 2 & 2 & 3 & b_2 \\ 4 & 4 & 8 & b_3 \end{array}\right] \Rightarrow \begin{array}{c} \\ \\ R_3 - 4R_2 \end{array}\left[\begin{array}{ccc|c} 1 & 1 & 1 & b_1 \\ 0 & 0 & 1 & b_2 - 2b_1 \\ 0 & 0 & 4 & b_3 - 4b_1 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & b_1 \\ 0 & 0 & 1 & b_2 - b_1 \\ 0 & 0 & 0 & b_3 - 4b_2 + 4b_1 \end{array}\right]$$

as $R_3 - 4R_2 \Rightarrow b_3 - 4b_1 - 4(b_2 - 2b_1) = b_3 - 4b_2 + 4b_1$

This completes the forward elimination step. We see that (2) has a solution if and only if $b_3 - 4b_2 + 4b_1 = 0$ or $4b_1 - 4b_2 + b_3 = 0$. This means that the range of the operator T associated with the matrix A is the plane in $\mathbf{R}^3$ through the origin whose equation in x,y,z variables is

P: $4x - 4y + z = 0$ where $P \subseteq \mathbf{R}^3$. If $\vec{b} = \vec{b}_1 = [1,1,1]^T$ then $4(1) - 4(1) + 1 = 1 \neq 0$ and hence we have no solution. If $\vec{b} = \vec{b}_2 = [1,1,0]^T$ then $4(1) - 4(1) + 0 = 0$ so there exists a solution. In fact there exists an infinite number of solutions. The pivots are already ones. We finish the Gauss-Jordan process for $\vec{b} = \vec{b}_2 = [1,1,0]^T$. Our augmented matrix is

$$\begin{matrix} R_1 - R_2 \\ \\ \\ \end{matrix}\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow \begin{array}{rl} x + y & = 2 \\ z & = -1 \\ 0 & = 0 \end{array} \Rightarrow \begin{array}{l} x = 2 - y \\ z = -1 \\ \\ \end{array}$$

We can now write a (parametric) formula for the infinite number of solutions.

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

It will become clear why one wishes to writw the solution in this form after you learn to solve differential equations.

<u>CONVENIENT CHECKS</u>. There are convenient checks for the vectors that appear in the formula for the infinite number of solutions. Satisfying these checks does not guarantee a correct solution. However, failing these checks does guarantee an error.

Since we claim to have all solutions to $\underset{3x3}{A}\underset{3x1}{\vec{x}} = \underset{3x1}{\vec{b}}$, letting y = 0, we see that $\vec{x}_p = [2, 0, -1]^T$ should be a (particular) solution to $A\vec{x} = \vec{b}$. (There are an infinite number of solutions.)

$u + v + w = 1 \Rightarrow 2 + 0 - 1 = 1$
$2u + 2v + 5w = 1 \Rightarrow 2(2) + 2(0) + 5(-1) = -1$
$4u + 4v + 5w = 0 \Rightarrow 4(2) + 4(0) + 8(-1) = 0$

Hence this part of the formula for the solution checks. Careful analysis indicates that $\vec{x}_1 = [-1, 1, 0]^T$ should be a solution to the complementary (homogeneous) equation $A\vec{x} = \vec{0}$.

$u + v + w = 0 \Rightarrow -1 + 1 + 0 = 0$
$2u + 2v + 5w = 0 \Rightarrow 2(-1) + 2(1) + 5(0) = 0$
$4u + 4v + 8w = 0 \Rightarrow 4(-1) + 4(1) + 8(0) = 0$

Hence this part of the formula for the solution also checks.

EXERCISES on Possibilities for Linear Algebraic Equations

<u>EXERCISE #1</u>. True or False.

_____ 1. There exists three possibilities for the system $\underset{mxn}{A}\underset{nx1}{\vec{x}} = \underset{mx1}{\vec{b}}$.

_____ 2. One possibility for the system $\underset{mxn}{A}\underset{nx1}{\vec{x}} = \underset{mx1}{\vec{b}}$ is that there is no solution.

______ 3. One possibility for the system $\underset{m\times n}{A}\,\underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{b}}$ is that there exists exactly one solution.

______ 4. One possibility for the system $\underset{m\times n}{A}\,\underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{b}}$ is $\exists$ an infinite number of solutions.

<u>EXERCISE #2</u>. Solve

$$\begin{aligned} x + \quad y + z &= b_1 \\ 2x + 2y + 3z &= b_2 \\ 4x + 4y + 8z &= b_3 \end{aligned}$$

if $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ has the values 1) $\vec{b} = \vec{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ and 2) $\vec{b} = \vec{b}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$.

<u>EXERCISE #3</u>. Solve

$$\begin{aligned} x + \quad y + z &= b_1 \\ 2x + 2y + 3z &= b_2 \\ 4x + 4y + 4z &= b_3 \end{aligned}$$

if $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ has the values 1) $\vec{b} = \vec{b}_1 = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$ and 2) $\vec{b} = \vec{b}_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$.

Handout #2 WRITING SOLUTIONS OF LINEAR ALGEBRAIC EQUATIONS Professor Moseley

Consider the system of linear equations

$$\underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{b}} \tag{1}$$

Suppose we apply Gauss-Jordan elimination so that

$$\left[\underset{m\times n}{A}\middle|\underset{m\times 1}{\vec{b}}\right] \overset{GE}{\Rightarrow} \left[\underset{m\times n}{U}\middle|\underset{m\times 1}{\vec{c}}\right] \overset{GJE}{\Rightarrow} \left[\underset{m\times n}{R}\middle|\underset{m\times 1}{\vec{d}}\right]$$

where U is an upper triangular matrix obtained using Gauss Elimination, $\vec{c}$ is the result of applying these same operations to $\vec{b}$, R is the **<u>row- reduced-echelon</u> form** of A obtained using **Gauss - Jordan elimination**, and $\vec{d}$ is the result of applying these same operations to $\vec{b}$. This means that the matrix $\left[\underset{m\times n}{R}\middle|\underset{m\times 1}{\vec{d}}\right]$ has zeros above as well as below the pivots. Recall that there all three cases: 1) No solution, 2) Exactly one solution (a unique solution), and 3) An infinite number of solutions.

1) If there is a row in $\left[\underset{m\times n}{R}\middle|\underset{m\times 1}{\vec{d}}\right]$ where the entries in R are all zero and the entry in $\vec{d}$ is nonzero (this corresponds to an equation of the form $0 = d_i \neq 0$). Since this is not true, the system has no solution, and you conclude by writing: No Solution.

2) If there are $m - n$ rows of all zeros (including the elements in $\vec{d}$) and so n pivots (dependent variables) then each variable in $\vec{x}$ is determined uniquely and there is exactly one solution. Conclude by giving the solution explicitly in **both scalar and vector form**.

3) If there are an infinite number of solutions we proceed as follows:

a) Separate the variables (i.e. the columns) into two groups. Those corresponding to pivots are the **<u>dependent (basic) variables</u>**. Those not corresponding to pivots are the **<u>independent (free) variables</u>**.

b) Write the equations that correspond to the augmented matrix $[R \mid \vec{d}\,]$.

c) Solve these equations for the dependent variables in terms of the independent variables. (This is easy if GJE is used).

d) Replace the dependent variables in the vector $\vec{x} = [x_1, \dots, x_n]^T$ to obtain $\vec{x}$ in terms of the independent variables only.

e) Write the infinite number of solutions $\vec{x}$ in the paramentric form $\vec{x} = \vec{x}_p + \sum_{i=1}^{k} c_i \vec{x}_i$

where the c_i's are the independent variables. Thus we see that there is a solution for each set of paramenter values.

Alternately we see that $\vec{x}_p$ can be obtained by letting all the c_i's (i.e. the independent variables) be zero. If $\vec{b} = \vec{0}$ then $\vec{x}_p$ can be taken to be $\vec{0}$. The above procedure will in fact make it zero.) Hence the vectors $\vec{x}_I$ can be obtained by letting $\vec{b} = \vec{0}$ and letting each independent variable be one with the rest of the independent variables equal to zero.

<u>EXAMPLE#1 </u>. Recall that the solution process for

$$
\begin{aligned}
2x + y + z &= 1 && (1)\\
4x + y \quad\ &= -2 && (2)\\
-2x + 2y + z &= 7 && (3)
\end{aligned}
$$

using Gauss-Jordan elimination (right to left) on augmented matrices is

$$
\begin{array}{r} \\ R_2 - 2R_1 \\ R_3 + R_1 \end{array}
\left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 4 & 1 & 0 & -2 \\ -2 & 2 & 1 & 7 \end{array}\right]
\Rightarrow
\begin{array}{r} \\ \\ R_3 + 3R_2 \end{array}
\left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & -1 & -2 & -4 \\ 0 & 3 & 2 & 8 \end{array}\right]
\Rightarrow
$$

$$
\begin{array}{r} 1/2\,R_1 \\ -R_2 \\ -1/4\,R_3 \end{array}
\left[\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & -1 & -2 & -4 \\ 0 & 0 & -4 & -4 \end{array}\right]
\Rightarrow
\begin{array}{r} R_1 - 1/2R_3 \\ R_2 - 2R_3 \\ \\ \end{array}
\left[\begin{array}{ccc|c} 1 & 1/2 & 1/2 & 1/2 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 1 \end{array}\right]
$$

$$
\Rightarrow
\begin{array}{r} R_1 - 1/2R_2 \\ \\ \\ \end{array}
\left[\begin{array}{ccc|c} 1 & 1/2 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right]
\Rightarrow
\left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right]
\Rightarrow
$$

$$
\boxed{\begin{aligned} x &= -1 \\ y &= 2 \quad \text{Scalar form} \\ z &= 1 \end{aligned}}
\qquad
\vec{x} = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \ \text{Vector form.}
$$

<u>EXAMPLE#2 </u>. Recall that the solution process for

$$
\begin{array}{lrl}
\text{E1} & x + y + z &= b_1 \\
\text{E2} & 2x + 2y + 3z &= b_2 \qquad (2)\\
\text{E3} & 4x + 4y + 8z &= b_3
\end{array}
$$

if $\vec{b} = \vec{b}_1 = [1,1,1]^T$ using Gauss elimination (only the forward sweep is needed) on augmented matrices is

$$\begin{array}{c} \\ R_2-2R_1 \\ R_3-4R_1 \end{array}\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 2 & 3 & 1 \\ 4 & 4 & 8 & 1 \end{array}\right] \Rightarrow \begin{array}{c} \\ \\ R_3-4R_2 \end{array}\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 4 & -3 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{array}\right] \Rightarrow \begin{array}{r} x+y+z=1 \\ z=-1 \\ 0=1 \end{array}$$

The last equation (0=1) is not true so we see that the equations are **inconsistent** and that there is no solution. We write **NO SOLUTION**.

On the other hand, if $\vec{b}=\vec{b}_1 = [1,1,0]^T$ using Gauss-Jordan elimination on augmented matrices is

$$\begin{array}{c} \\ R_2-2R_1 \\ R_3-4R_1 \end{array}\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 2 & 3 & 1 \\ 4 & 4 & 8 & 0 \end{array}\right] \Rightarrow \begin{array}{c} \\ \\ R_3-4R_2 \end{array}\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 4 & -3 \end{array}\right] \Rightarrow \begin{array}{c} R_1-R_2 \\ \\ \\ \end{array}\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right] \Rightarrow \begin{array}{c} R_1-R_2 \\ \\ \\ \end{array}\left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\Rightarrow \begin{array}{r} x+y\qquad=2 \\ z=-1 \\ 0=0 \end{array} \Rightarrow \begin{array}{l} x=2-y \\ z=-1 \end{array}$$

We can now write a (parametric) formula for the infinite number of solutions.

$$\begin{array}{l} x=2-y \\ z=-1 \end{array} = \begin{bmatrix} 2-y \\ y \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} + y\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{Vector Form} \qquad \begin{array}{l} x=1+y \\ (y=y) \quad \text{Scalar Form} \\ z=-1 \end{array}$$

EXERCISES on Writing Solutions of Linear Algebraic Equations

EXERCISE #1. Suppose the Gauss process yields

$$\begin{aligned} x + \quad y + z &= b_1 \\ 2y + 3z &= b_2 \\ 8z &= b_3 \end{aligned}$$

If $\vec{b} = [b_1, b_2, b_3]^T = \vec{b}_1 = [1,1,1]^T$, give the solution and the solution set. What is the difference in these two questions. How many solutions are there? The solution should be written as ______.
If $\vec{b} = [b_1, b_2, b_3]^T = \vec{b}_1 = [0,0,0]^T$, give the solution and the solution set. What is the difference in these two questions. How many solutions are there? The solution should be written as ______.

EXERCISE #2. Suppose the Gauss-Jordan process yields $\left[\begin{array}{ccc|c} 1 & 0 & 0 & b_1 - b_2 \\ 0 & 1 & 0 & b_2 - b_1 \\ 0 & 0 & 0 & b_3 - b_2 - b_1 \end{array}\right]$

If $\vec{b} = [b_1, b_2, b_3]^T = \vec{b}_1 = [1,1,1]^T$, give the solution and the solution set. What is the difference in these two questions. How many solutions are there? The solution should be written as ______.
If $\vec{b} = [b_1, b_2, b_3]^T = \vec{b}_1 = [1,-1,0]^T$, give the solution and the solution set. What is the difference in these two questions. How many solutions are there? The solution should be written as _____.

DEFINITION #1. An mxm matrix is said to be an elementary matrix if it can be obtained from the mxm identity matrix $\underset{mxm}{I}$ by means of a single ERO.

THEOREM #1. Let e(A) denote the matrix obtained when the ERO e is applied to the matrix A. Now let E = e(I). Then for any mxn matrix A, we have e(A) = EI. That is, the effect of an ERO e on a matrix A can be obtain by multiplying A by E where E = e(I). Thus E is called the **matrix representation** of e.

THEOREM#2. All elementary matrices are invertible. If E is an elementary matrix that represents the ERO e with ERO inverse e^{-1}, then E^{-1} is just the matrix representation of e^{-1}.

Suppose we apply Gauss-Jordan elimination to $\underset{mxn}{A}\,\underset{nx1}{\vec{x}} = \underset{mx1}{\vec{b}}$ so that

$$\left[\underset{mxn}{A}\middle|\underset{mx1}{\vec{b}}\right] \overset{GJE}{\Rightarrow} \left[\underset{mxn}{R}\middle|\underset{mx1}{\vec{d}}\right]$$

via the sequence of ERO's $e_1, e_2, ..., e_n$ and $E_i = e_i(I)$ as well as $E_i^{-1} = e_i^{-1}(I)$ for i=1, 2, 3, ..., n. then

$$E_n\, E_{n-1} \cdots E_2\, E_1\, A\,\vec{x} = R\,\vec{x} = E_n\, E_{n-1} \cdots E_2\, E_1\,\vec{b} = \vec{d}$$

That is, we can solve the vector equation by repeatedly multiplying the vector (or matrix) equation $\underset{mxn}{A}\,\underset{nx1}{\vec{x}} = \underset{mx1}{\vec{b}}$ by matrix representations of the appropriate ERO's until we obtain $R\,\vec{x} = \vec{d}$. If we are in the unique solution case, then $n \le m$, the first n row of R are the nxn identity matrix I and we obtain $\vec{x} = \vec{d}_r$ where $\vec{d}_r$ is the nx1 vector containing the first n components of $\vec{d}$ whose remaining m − n components are all zeros. If not, we must examine R and $\vec{d}$ more closely to determine which case we are in. The main point is that instead of working with the scalar equations or the augmented matrices, we may solve the vector (or matrix) equation $\underset{mxn}{A}\,\underset{nx1}{\vec{x}} = \underset{mx1}{\vec{b}}$ by multiplying it successively by (invertible) elementary matrices to obtain an equivalent form of the vector equation where the unique solution is readily available, we know immediately that there is no solution, or we may easily find a parametric form for all of the infinite number of solutions.

EXERCISES on Elementary Matrices and Solving the Vector Equations

<u>EXERCISE #1</u>. True or False

_____ 1. An mxm matrix is said to be an elementary matrix if it can be obtained from the mxm identity matrix $\underset{mxm}{I}$ by means of a single ERO

_____ 2. If e(A) denotes an ERO applied to the matrix A and let E = e(I), then for any mxn matrix A, we have e(A) = EI.

_____ 3. All elementary matrices are invertible.

_____ 4. If E is an elementary matrix that represents the ERO e with ERO inverse e^{-1}, then E^{-1} is just the matrix representation of e^{-1}.

A SERIES OF CLASS NOTES TO INTRODUCE LINEAR AND NONLINEAR PROBLEMS
TO ENGINEERS, SCIENTISTS, AND APPLIED MATHEMATICIANS

LINEAR CLASS NOTES:
A COLLECTION OF HANDOUTS FOR
REVIEW AND PREVIEW
OF LINEAR THEORY
INCLUDING FUNDAMENTALS OF
LINEAR ALGEBRA

CHAPTER 5

Linear Operators,

Span, Linear Independence,

Basis Sets, and Dimension

In this handout, we preview **linear operator** theory. The most important examples of linear operators are differential and integral operators and operators defined by matrix multiplication. These arise in many applications. **Lumped parameter systems** (e.g., linear circuits and mass spring systems) give rise to **discrete operators** defined on finite dimensional vector spaces (e.g., $\mathbf{R}^n$). Differential and integral equations (e.g., Maxwell's equations and the Navier-Stokes equation) are used to model **distributed (continuum) systems** and require infinite dimensional vector spaces. These give rise to differential and integral operators on **function spaces**.

Even without covering any topics in differential equations, your background in calculus should be sufficient to see how discrete and continuous operators are connected as linear operators on a vector space.

A function or map T from one vector space V to another vector space W is often call an **operator**. If we wish to think geometrically (e.g., if V and W are R^2 or R^3) rather than algebraically we might call T a transformation.

DEFINITION 1. Let V and W be vector spaces over the same field **K**. An operator $T:V \to W$ is said to be **linear** if $\forall \vec{x}, \vec{y} \in V$ and scalars $\alpha,\beta\in\mathbf{K}$, it is true that

$$T(\alpha \vec{x} + \beta \vec{y}) = \alpha T(\vec{x}) + \beta T(\vec{y}). \tag{1}$$

THEOREM 1. Let V and W be vector spaces over the same field **K**. An operator $T: V \to W$ is linear if and only if the following two properties hold:

i) $\vec{x}, \vec{y} \; \varepsilon \; V$ implies $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y})$ (2)

ii) $\alpha\in\mathbf{K}$ and $\vec{x} \; \varepsilon \; V$ implies $T(\alpha\vec{x}) = \alpha T(\vec{x})$. (3)

EXAMPLE 1 Let the **operator** $T:\mathbf{R}^n \to \mathbf{R}^m$ be defined by **matrix multiplication** of the column vector $\vec{x}$ by the $m\times n$ matrix A; that is, let

$$T(\vec{x}) =_{df} \underset{mxn}{A} \, \underset{nx1}{\vec{x}} \tag{4}$$

where $\underset{mxn}{\vec{x}} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \in\mathbf{R}^n$ and $\underset{mxn}{A} = \begin{bmatrix} a_{11} & a_{12} & \cdot & \cdot & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & \cdot & \cdot & a_{2n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{m1} & a_{m21} & \cdot & \cdot & \cdot & a_{mn} \end{bmatrix} \in \mathbf{R}^{m\times n}$. Then T is a linear operator.

EXAMPLE 2 Let I = (a,b). The operator $D:C^1(I,\mathbf{R}) \to C(I,\mathbf{R})$ defined by

$$D(f) =_{df} \frac{df}{dx} \tag{5}$$

where $f \varepsilon C^1(I,\mathbf{R}) = \{f:I \to \mathbf{R}: \frac{df}{dx}$ exists and is continuous on I} and
$C(I, \mathbf{R}) = \{f:I \to \mathbf{R}$:f is continuous on I}. Then D is a linear operator. We may restrict D to $A(I, \mathbf{R}) = \{f:I \to \mathbf{R}$:f is analytic on I} so that $D:A(I, \mathbf{R}) \to A(I, \mathbf{R})$ maps a vector space back to itself.

DEFINITION #2. Let T:V→W be a mapping from a set V to a set W. The set
R(T) = {y∈ W: there exists an x∈V such that y = T(x) } is called the **range** of T. If W has an additive structure (i.e., a binary operation which we call addition) with an additive identity which we call 0, then the set N(T) = {x∈V: T(x) = 0} is called the **null set** of T (or nullity of T).

If T is a linear operator from a vector space V to another vector space W, we can say more.

THEOREM #2 Let T:V→W be a linear operator from a vector space V to a vector space W. The range of T, $R(T) = \{ \vec{y} \in W$: there exists an $\vec{x} \in V$ such that $\vec{y} = T(\vec{x})$ }, is a subspace of W and the null set of T, $N(T) = \{ \vec{x} \in V: T(\vec{x}) = 0\}$ is a subspace of V.

We rename these sets.

DEFINITION #3. Let T:V→W be a linear operator from a vector space V to another vector space W. The set $R(T) = \{ \vec{y} \in W$: there exists an $\vec{x} \in V$ such that $\vec{y} = T(\vec{x})$ } is called the **range space** of T and the set $N(T) = \{ \vec{x} \in V: T(\vec{x}) = \vec{0} \}$ is called the **null space** of T. .

EXERCISES on Linear Operator Theory

<u>EXERCISE #1</u>. True or False.

_____ 1. A linear circuit is an example of a lumped parameter system.

_____ 2. Amass/spring system is an example of a lumped parameter system.

_____ 3. A function or map T from one vector space V to another vector space W is often call an operator.

_____ 4. An operator $T:V \to W$ is said to be linear if $\forall \vec{x}, \vec{y} \in V$ and scalars $\alpha,\beta \in \mathbf{K}$, it is true that $T(\alpha \vec{x} + \beta \vec{y}) = \alpha T(\vec{x}) + \beta T(\vec{y})$.

_____ 5. An operator $T: V \to W$ is linear if and only if the following two properties hold: i) $\vec{x}, \vec{y}\ \varepsilon\ V$ implies $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y})$ and ii) $\alpha \in \mathbf{K}$ and $\vec{x}\ \varepsilon\ V$ implies

$T(\alpha \vec{x}) = \alpha T(\vec{x})$.

_____ 6. The operator $T:\mathbf{R}^n \to \mathbf{R}^m$ defined by $T(\vec{x}) = \underset{m\mathrm{x}n}{A}\ \underset{n\mathrm{x}1}{\vec{x}}$ is a linear operator.

_____ 7. The operator $D:C^1(a,b) \to C(a,b)$ be defined by $D(f) =_{df} \dfrac{df}{dx}$ is a linear operator.

_____ 8. $C(a,b) = \{f:(ab) \to \mathbf{R}$:f is continuous$\}$.

_____ 9. $C^1(a,b) = \{f:(a,b) \to \mathbf{R}$: $\dfrac{df}{dx}$ exists and is continuous$\}$.

_____ 10. If $T:V\to W$, then the set $R(T) = \{y\in W$: there exists an $x\in V$ such that $y = T(x)\}$ is called the range of T.

_____ 11. If $T:V\to W$ and W has an additive structure (i.e., a binary operation which we call addition) with an additive identity which we call 0, then the set $N(T) = \{x\in V: T(x) = 0\}$ is called the null set of T (or nullity of T).

_____ 12. If $T:V\to W$ is a linear operator from a vector space V to a vector space W, then the range of T, $R(T) = \{\vec{y} \in W$: there exists an $\vec{x} \in V$ such that $\vec{y} = T(\vec{x})\}$, is a subspace of W

_____ 14.If $T:V\to W$ is a linear operator from a vector space V to a vector space W, then the null set of T, $N(T) = \{\vec{x} \in V: T(\vec{x}) = 0\}$ is a subspace of V.

_____ 15. If $T:V\to W$ is a linear operator from a vector space V to another vector space W, then the set $R(T) = \{\vec{y} \in W$: there exists an $\vec{x} \in V$ such that $\vec{y} = T(\vec{x})\}$ is called the range space of T.

_____ 16. If $T:V\to W$ is a linear operator from a vector space V to another vector space W, then the set $N(T) = \{\vec{x} \in V: T(\vec{x}) = \vec{0}\}$ is called the null space of T.

Handout # 2 SPANNING SETS Professor Moseley

DEFINITION #1. If $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_n$ are vectors in a vector space V and $\alpha_1,\alpha_2,\cdots,\alpha_n$ are scalars, then

$$\alpha, \ \vec{x}_1 + \alpha_2 \ \vec{x}_2 + \cdots + \alpha_n \ \vec{x}_n = \sum_{i=1}^{n} \alpha_i \ \vec{x}_1$$

is called a **linear <u>combination</u>** of the vectors. (It is important to note that a linear combination allows only a finite number of vectors.)

EXAMPLE #1. Consider the system of linear algebraic equations:

$$\begin{aligned} 2x + y + z &= 1 \\ 4x + y \quad\ &= -3 \\ x - 2y - z &= 0 \quad . \end{aligned}$$

This set of **scalar equations** can be written as the **vector equation**

$$x \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} + y \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} + z \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ 0 \end{bmatrix}$$

where the left hand side (LHS) is a linear combination of the (column vectors whose components come from the) columns of the coefficient matrix

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

If we generalize Example #1, we obtain:

THEOREM #1. The general system

$$\underset{nxn}{A} \ \underset{nx1}{\vec{x}} = \underset{nx1}{\vec{b}} \qquad (1)$$

has a solution if and only if the vector $\vec{b}$ can be written as a linear combination of the columns of A. That is, $\vec{b}$ is in the range space of $T(\vec{x}) = A\vec{x}$ if and only if $\vec{b}$ can be written as a linear combination of the (column vectors whose components come from the) columns of the coefficient matrix. A.

DEFINITION #2. Let S be a (finite) subset of a subspace W of a vector space V. If every vector in W can be written as a linear combination of (a finite number of) vectors in S, then S is said to **span** W or to form a **spanning set** for W. On the other hand, if S is any (finite) set of vectors, the **span** of S, written Span(S), is the set of all possible linear combinations of (a finite number of) vectors in S.

THEOREM #2. For any (finite) subset S of a vector space V, Span (S) is a subspace of V.

THEOREM #3. If (a finite set) S is a spanning set for W, then Span S = W.

EXAMPLES. Consider the following subsets of $\mathbf{R}^3$.

1. $S = \{[1,0,0]^T\} = \{\hat{\imath}\}$. Then Span (S) $= \{x\,\hat{\imath}: x \in \mathbf{R}\}$ = x - axis.
2. $S = \{[1,0,0]^T, [0,1,0]^T\} = \{\hat{\imath}, j\}$. Then Span (S) $= \{x\,\hat{\imath}+y\,\hat{\jmath}: x, y \in \mathbf{R}\}$ = the xy plane.
3. $S = \{[1,0,0]^T, [1,1,0]^T\} = \{\hat{\imath}, \hat{\imath}+\hat{\jmath}\}$. Then Span (S) = xy plane.
4. $S = \{[1,0,0]^T, [0,1,0]^T, [1,1,0]^T\} = \{\hat{\imath}, \hat{\jmath}, \hat{\imath}+\hat{\jmath}\}$. Then Span (S) = xy plane.

DEFINITION #3. For any matrix $\underset{m\times n}{A}$, the span of the set of column vectors is the **column space** of A. The usual notation for the column space is R(A) and sometimes R_A. Normally we will use R(A). The reason for this notation is that when we think of the operator $T(\vec{x})$ defined by (matrix) multiplication of the matrix $\underset{m\times n}{A}$ by the column vector $\vec{x} = [x_1,\cdots,x_n]^T$, we see that T maps vectors $\vec{x} \in \mathbf{R}^n$ into vectors $\vec{y} = A\vec{x} \in \mathbf{R}^m$, the column space R(A) is seen to be the **range** (space) of T.

COMMENT. Consider the general system (1) above. Theorem #1 can now be rephrased to say that (1) has a solution if and only if $\vec{b}$ is in the column space of A (i.e. in the range of the operator T).

DEFINITION 4. For any matrix $\underset{m\times n}{A}$, the span of the set of row vectors is the **row space** of A. The usual notation for the row space is $R(A^T)$ and sometimes R_{A^T} since the row space of A is the column space of A^T. Normally we will use $R(A^T)$.

If we think of the operator $T^T(\vec{x})$ defined by (matrix) multiplication of the matrix $\underset{m\times n}{A}^T$ by the column vector $\vec{y} = [y_1,\cdots,y_m]^T$, we see that A^T maps vectors $\vec{y} \in \mathbf{R}^m$ into vectors $\vec{x} = A^T\vec{y} \in \mathbf{R}^n$, the row space $R(A^T)$ is seen to be the **range** (space) of A^T.

Recall that all of the coefficient matrices for the associated linear systems of algebraic equations obtained in the process of doing Gauss elimination are row equivalent. Hence they all have the same row space. However, they do not have the same column space.

EXERCISES on Spanning Sets

EXERCISE #1. True or False.

_____ 1. If $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_n$ are vectors in a vector space V and $\alpha_1, \alpha_2, \cdots, \alpha_n$ are scalars, then

$$\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \cdots + \alpha_n \vec{x}_n = \sum_{i=1}^{n} \alpha_i \vec{x}_1$$ is a linear combination of the vectors

_____ 2. A linear combination allows only a finite number of vectors.

_____ 3. The linear algebraic equations: $2x + y + z = 1$, $4x + y = -3$, $x - 2y - z = 0$, is a system of scalar equations.

_____ 4. The coefficient matrix for the system of linear algebraic equations: $2x + y + z = 1$, $4x + y = -3$, $x - 2y - z = 0$ is $A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 1 & 0 \\ 1 & -2 & -1 \end{bmatrix}$.

_____ 5. The linear algebraic equations: $2x + y + z = 1$, $4x + y = -3$, $x - 2y - z = 0$ can be written as the vector equation $x\begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} + y\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} + z\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ 0 \end{bmatrix}$.

_____ 5. The left hand side (LHS) of the vector equation $x\begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} + y\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} + z\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ $= \begin{bmatrix} 1 \\ -3 \\ 0 \end{bmatrix}$ is a linear combination of the column vectors whose components come from the columns of the coefficient matrix $A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 1 & 0 \\ 1 & -2 & -1 \end{bmatrix}$.

_____ 6. The general system $\underset{n\times n}{A}\ \underset{n\times 1}{\vec{x}} = \underset{n\times 1}{\vec{b}}$ has a solution if and only if the vector $\vec{b}$ can be written as a linear combination of the columns of A.

_____ 7. $\vec{b}$ is in the range space of T($\vec{x}$) = A $\vec{x}$ if and only if $\vec{b}$ can be written as a linear combination of the column vectors whose components come from the columns of the coefficient matrix A.

_____ 8. If S be a finite subset of a subspace W of a vector space V and every vector in W can be written as a linear combination of a finite number of vectors in S, then S is said to span W or to form a spanning <u>set</u> for W.

_____ 9. If S be a finite subset of a subspace W of a vector space V, then the span of S, written Span(S), is the set of all possible linear combinations of a finite number of vectors in S.

_____ 10. For any finite subset S of a vector space V, Span (S) is a subspace of V.

_____ 11. If S be a finite subset of a subspace W of a vector space V and S is a spanning set for W, then Span S = W.

_____ 12. For any matrix $\underset{m\times n}{A}$, the span of the set of column vectors is the column space of A.

_____ 13. The usual notation for the column space is R(A) and sometimes R_A.

_____ 14. The reason that R(A) is the column space of A is that when we think of the operator T($\vec{x}$) defined by matrix multiplication of the matrix $\underset{m\times n}{A}$ by the column vector $\vec{x} = [x_1,\cdots,x_n]^T$, we see that T maps vectors $\vec{x} \in \mathbf{R}^n$ into vectors $\vec{y} = A\vec{x} \in \mathbf{R}^m$ which is the column space R(A)

_____ 14. The column space R(A) is the range space of T.

_____ 15. $\underset{n\times n}{A}\ \underset{n\times 1}{\vec{x}} = \underset{n\times 1}{\vec{b}}$ has a solution if and only if $\vec{b}$ is in the range of the operator T($\vec{x}$) = A $\vec{x}$.

_____ 16. $\underset{n\times n}{A}\ \underset{n\times 1}{\vec{x}} = \underset{n\times 1}{\vec{b}}$ has a solution if and only if $\vec{b}$ is in the column space of A.

_____ 17. For any matrix $\underset{m\times n}{A}$, the span of the set of row vectors is the row space of A.

_____ 18. The usual notation for the row space is $R(A^T)$ and sometimes R_{A^T} since the row space of A is the column space of A^T.

_____ 19. If we think of the operator $T^T(\vec{x})$ defined by (matrix) multiplication of the matrix $\underset{m\times n}{A}^{T}$ by the column vector $\vec{y} = [y_1,\cdots,y_m]^T$, we see that A^T maps vectors $\vec{y} \in \mathbf{R}^m$ into vectors $\vec{x} = A^T\vec{y} \in \mathbf{R}^n$, the row space $R(A^T)$ is seen to be the range space of A^T.

DEFINITION #1. Let V be a vector space. A finite set of vectors $S = \{\vec{x}_1, \ldots, \vec{x}_k\} \subseteq V$ is **linearly independent** (ℓ.i.) if the only set of scalars $c_1, c_2, \ldots, c_k$ which satisfy the (homogeneous) vector equation

$$c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_k\vec{x}_k = \vec{0} \tag{1}$$

is $c_1 = c_2 = \cdots = c_n = 0$; that is, (1) has only the trivial solution. If there is a set of scalars not all zero satisfying (1) then S is **linearly dependent** (ℓ.d.).

It is common practice to describe the vectors (rather than the set of vectors) as being linearly independent or linearly dependent. Although this is technically incorrect, it is wide spread, and hence we accept this terminology. Since we may consider (1) as a linear homogeneous equation with unknown vector $[\, c_1, c_2, \ldots, c_n] \in \mathbf{R}^n$, if (1) has one nontrivial solution, then it in fact has an infinite number of nontrivial solutions. As part of the standard procedure for showing that a set is **linearly dependent** directly using the definition (DUD) you must exhibit one (and only one) such non trivial solution. To show that a set is **linearly independent** directly using the definition (DUD) you must show that the only solution of (1) is the trivial solution (i.e. that all the c_i's must be zero). Before looking at the application of this definition to column vectors, we state four theorems.

THEOREM #1. Let V be a vector space. If a finite set of vectors $S = \{\vec{x}_1, \ldots, \vec{x}_k\} \subseteq V$ contains the zero vector, then S is linearly dependent.

THEOREM #2. Let V be a vector space. If $\vec{x} \neq \vec{0}$, then $S = \{\vec{x}\} \subseteq V$ is linearly independent.

Proof. To show S is linearly independent we must show that $c_1\vec{x} = \vec{0}$ implies that $c_1 = 0$. But by the zero product theorem, if $c_1\vec{x} = \vec{0}$ is true then $c_1 = 0$ or $\vec{x} = \vec{0}$. But by hypothesis $\vec{x} \neq \vec{0}$. Hence $c_1\vec{x} = \vec{0}$ implies $c_1 = 0$. Hence $S = \{\vec{x}\}$ where $\vec{x} \neq \vec{0}$ is linearly independent.

Q.E.D.

THEOREM #3. Let V be a vector space and $S = \{\vec{x}, \vec{y}\} \subseteq V$. If either $\vec{x}$ or $\vec{y}$ is the zero vector, then S is linearly dependent.

THEOREM #4. Let V be a vector space and $S = \{\vec{x}, \vec{y}\} \subseteq V$ where $\vec{x}$ and $\vec{y}$ are nonzero vectors. Then S is linearly dependent if and only if one vector is a scalar multiple of the other.

Although the definition is stated for an abstract vector space and hence applies to any vector space and we have stated some theorems in this abstract setting, in this section we focus on column vectors in $\mathbf{R}^n$ (or $\mathbf{C}^n$ or $\mathbf{K}^n$). Since we now know how to solve a system of linear algebraic equations, using this procedure, we can develop a "procedure" to show that a finite set in $\mathbf{R}^n$ (or $\mathbf{C}^n$ or $\mathbf{K}^n$) is linearly independent. We also show how to give sufficient conditions to show that a finite set in $\mathbf{R}^n$ (or $\mathbf{C}^n$ or $\mathbf{K}^n$) is linearly dependent.

PROCEDURE. To determine if a set $S = \{\vec{x}_1,\ldots,\vec{x}_k\} \subseteq V$ is linearly independent or linearly dependent, we first write down the equation (1) and try to solve. If we can show that the only solution is the trivial solution $c_1 = c_2 = \cdots = c_n = 0$, then we have shown **directly using the definition** (DUD) of linear independence that S is linearly independent. On the other hand (OTOH), if we can exhibit a nontrivial solution, then we have shown **directly using the definition** of linear dependence that S is linearly dependent. We might recall that the linear theory assures us that if there is one nontrivial solution, that there are an infinite number of nontrivial solutions. However, to show that S is linearly dependent directly using the definition, it is not necessary (or desirable) to find all of the nontrivial solutions. Although you could argue that once you are convinced (using some theorem) that there are an infinite number of solutions, then we do not need to exhibit one, this will not be considered to be a proof directly using the definition (as it requires a theorem). Thus to prove that S is linearly dependent directly using the definition of linear dependence, **you must exhibit one nontrivial solution**. This will help you to better understand the concepts of linear independence and linear dependence. We apply these "procedures" to finite sets in $\mathbf{R}^n$ (or $\mathbf{C}^n$). For $S = \{\vec{x}_1,\ldots,\vec{x}_k\} \subseteq \mathbf{R}^n$ (or $\mathbf{C}^n$) (1) becomes a system of n equations (since we are in $\mathbf{R}^n$ (or $\mathbf{C}^n$)) in k unknowns (since we have k vectors). (This can be confusing when applying general theorems about m equations in n unknowns. However, this should not be a problem when using DUD on specific problems.)

EXAMPLE #1. Determine (using DUD) if $S = \left\{ \begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}2\\3\end{bmatrix}, \begin{bmatrix}1\\2\end{bmatrix} \right\}$ is linearly independent.

Solution. (This is not a yes-no question and a proof is required). Assume

$$c_1\begin{bmatrix}1\\1\end{bmatrix} + c_2\begin{bmatrix}2\\3\end{bmatrix} + c_3\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix} \qquad (2)$$

and (try to) solve. The vector equation (2) is equivalent to the two scalar equations (since we are in $\mathbf{R}^2$) in three unknowns (since S has three vectors).

$$\begin{aligned} c_1 + 2c_2 + c_3 &= 0 \\ c_1 + 3c_2 + 2c_3 &= 0 \end{aligned}$$

(3)

Simple systems can often be solved using ad hoc (for this case only) procedures. But for complicated systems we might wish to write this system in the form $A\vec{c} = \vec{0}$ and use Gauss elimination (on a computer). Note that when we reduce A we do not have to augment. Why?

For this example $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \end{bmatrix}$ and $\vec{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$. Since the system is homogeneous we can solve by reducing A without augmenting (Why?).

$$R_1 - R_2 \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \Rightarrow \begin{matrix} c_1 + 2c_2 = c_3 = 0 \\ c_2 + c_3 = 0 \end{matrix} \Rightarrow \begin{matrix} c_1 = -2c_2 - c_3 = 2c_3 - c_3 = c_3 \\ c_2 = -c_3 \end{matrix}$$

Hence the general solution of (2) is $\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} c_3 \\ -c_3 \\ c_3 \end{bmatrix} = c_3 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$.

Hence there are an infinite number of solutions. They are the vectors in the subspace $W = \{ \vec{x} \in R^3 : \vec{x} = c\,[2, -1, 1]^T \text{ with } c \in \mathbf{R}\}$. Since there is a nontrivial solution, S is linearly dependent. However, we <u>must</u> exhibit one nontrivial solution which we do by choosing $c_1=1$, $c_2 = -1$, and $c_3 = 1$. Hence we have

$$(1)\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (-1)\begin{bmatrix} 2 \\ 3 \end{bmatrix} + (1)\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{4}$$

Since we have exhibited a nontrivial linear combination of the vectors in S, (4) alone proves that S is a linearly dependent set in the vector space $\mathbf{R}^2$ QED

It may be easy to guess a nontrivial solution (if one exists). We call this method the Clever Ad Hoc (CAH) method. You may have noted that the first and third vectors in the previous example add together to give the second vector. Hence the coefficients c_1, c_2, and c_3 could have been easily guessed.

<u>EXAMPLE #2</u>. Determine if $S = \left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \right\}$ is linearly independent.

<u>Solution</u>. (Again this is not a yes-no question and a proof is required) Note that since

$3\begin{bmatrix}1\\2\\-1\end{bmatrix}=\begin{bmatrix}3\\6\\-3\end{bmatrix}$ we have

$$(3)\begin{bmatrix}1\\2\\-1\end{bmatrix}+(-1)\begin{bmatrix}3\\6\\-3\end{bmatrix}+(0)\begin{bmatrix}4\\5\\6\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}. \qquad (5)$$

Since we have exhibited a nontrivial linear combination of the vectors in S, (5) alone proves that S is a linearly dependent set in the vector space $\mathbf{R}^3$ Hence S is linearly dependent. Q.E.D.

We give a final example.

<u>EXAMPLE #3.</u> Determine if $S=\left\{\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}1\\2\end{bmatrix}\right\}$ is linearly independent.

<u>Solution</u>. Since $\begin{bmatrix}1\\1\end{bmatrix}$ is not a multiple of $\begin{bmatrix}1\\2\end{bmatrix}$ we assume

$$c_1\begin{bmatrix}1\\1\end{bmatrix}+c_2\begin{bmatrix}1\\2\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}\Rightarrow \begin{matrix}c_1+c_2=0\\c_1+2c_2=0\end{matrix}\Rightarrow A\vec{c}=\vec{0}$$

where

$$A=\begin{bmatrix}1&1\\1&2\end{bmatrix} \text{ and } \vec{c}=\begin{bmatrix}c_1\\c_2\end{bmatrix}. \text{ Hence}$$

$$\begin{matrix}\\R_1-R_2\end{matrix}\begin{bmatrix}1&1\\1&2\end{bmatrix}\Rightarrow\begin{bmatrix}1&1\\0&1\end{bmatrix}\Rightarrow\begin{matrix}c_1+c_2=0\\c_2=0\end{matrix}\Rightarrow c_1=c_2=c_2=0$$

Since we have proved that the trivial linear combination where $c_1 = c_2 = 0$ is the only linear combination of the vectors in S that gives the zero vector in $\mathbf{R}^2$ (i.e., we have proved that S is a linearly independent set.

EXERCISES on Linear Independence (of Column Vectors)

EXERCISE #1. True or False.

_____ 1. If V be a vector space and $S = \{\vec{x}_1,\ldots,\vec{x}_k\} \subseteq V$, then S is linearly independent (ℓ.i.) if the only set of scalars c_1, c_2, ..., c_k which satisfy the (homogeneous) vector equation $c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_k\vec{x}_k = \vec{0}$ is $c_1 = c_2 = \cdots = c_n = 0$.

_____ 2. If V be a vector space and $S = \{\vec{x}_1,\ldots,\vec{x}_k\} \subseteq V$, then S is linearly independent (ℓ.i.) if the only set of scalars c_1, c_2, ..., c_k which satisfy the (homogeneous) vector equation $c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_k\vec{x}_k = \vec{0}$ is the trivial solution.

_____ 3. If V be a vector space and $S = \{\vec{x}_1,\ldots,\vec{x}_k\} \subseteq V$, then S is linearly dependent (ℓ.d.) if there is a set of scalars not all zero satisfying $c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_k\vec{x}_k = \vec{0}$.

_____ 4. If V is a vector space and $S = \{\vec{x}_1,\ldots,\vec{x}_k\} \subseteq V$ contains the zero vector, then S is linearly dependent.

_____ 5. If V be a vector space, $\vec{x} \neq \vec{0}$ and $S = \{\vec{x}\} \subseteq V$, then S is linearly independent.

_____ 6. If $c_1\vec{x} = \vec{0}$ is true then by the zero product theorem either $c_1 = 0$ or $\vec{x} = \vec{0}$. _____

_____ 7. If V is a vector space and $S = \{\vec{x}, \vec{y}\} \subseteq V$ and either $\vec{x}$ or $\vec{y}$ is the zero vector, then S is linearly dependent.

_____ 8. If V is a vector space and $S = \{\vec{x}, \vec{y}\} \subseteq V$ where $\vec{x}$ and $\vec{y}$ are nonzero vectors, then S is linearly dependent if and only if one vector is a scalar multiple of the other.

EXERCISE #2. Use the procedure given above to determine (using DUD) if $S = \left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}4\\3\end{bmatrix}, \begin{bmatrix}2\\2\end{bmatrix}\right\}$ is linearly dependent or linearly independent or neither. Thus you must explain completely.

EXERCISE #3. Use the procedure given above to determine (using DUD) if $S = \left\{\begin{bmatrix}1\\2\\-1\end{bmatrix}, \begin{bmatrix}4\\8\\-4\end{bmatrix}, \begin{bmatrix}4\\5\\6\end{bmatrix}\right\}$ is linearly dependent or linearly independent or neither. Thus you must explain completely.

EXERCISE #4. Use the procedure given above to determine (using DUD) if $S = \left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\2\end{bmatrix}\right\}$ is linearly dependent or linearly independent or neither. Thus you must explain completely.

EXERCISE #5. Prove Theorem #1.

EXERCISE #6. Prove Theorem #3.

EXERCISE #7. Prove Theorem #4.

DEFINITION #1. Let B = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_k$} ⊆ W ⊆ V where W is a subspace of the vector space V. Then B is a **basis** of W if

i) B is linearly independent

ii) B spans W (i.e. Span B = W)

To prove that a set B is a basis (or basis set or base) for W we must show both i) and ii). We already have a method to show that a set is **linearly independent**. To use DUD consider the vector equation

$$c_1 \vec{x}_1 + c_2 \vec{x}_2 + \cdots + c_k \vec{x}_k = \vec{0} \qquad (1)$$

in the unknown variables c_1, c_2, ...,c_k and show that the trivial solution $c_1 = c_2 \cdots = c_k = 0$ is the only solution of (1). To show that B is a **spanning set** using DUD we must show that an arbitrary vector $\vec{b}$ ∈ W can be written as a linear combination of the vectors in B; that is we must show that the vector equation

$$c_1 \vec{x}_1 + c_2 \vec{x}_2 + \cdots + c_k \vec{x}_k = \vec{b} \qquad (2)$$

in the unknown variables c_1, c_2, ...,c_k always has at least one solution.

EXAMPLE (THEOREM) #1. Show that B = { $[1,0,0]^T$, $[1,1,0]^T$ } is a basis for W = { $[x, y, 0]^T$: x,y ∈ **R** }.

Solution. (proof) To show linear independence we solve

$$c_1\begin{bmatrix}1\\0\\0\end{bmatrix}+c_2\begin{bmatrix}1\\1\\0\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix} \quad \text{or} \quad \begin{aligned} c_1 + c_2 &= 0\\ c_2 &= 0\\ 0 &= 0\end{aligned} \quad \text{to obtain } c_1 = c_2 = 0$$

so that B is linearly independent.

ii) To show B spans W we let $\vec{x}$ = [x, y, 0] ∈ W (i.e., an arbitrary vector in W) and solve

$$c_1\begin{bmatrix}1\\0\\0\end{bmatrix}+c_2\begin{bmatrix}1\\1\\0\end{bmatrix}=\begin{bmatrix}x\\y\\0\end{bmatrix} \quad c_1\begin{bmatrix}1\\0\\0\end{bmatrix}+c_2\begin{bmatrix}1\\1\\0\end{bmatrix}=\begin{bmatrix}x\\y\\0\end{bmatrix} \quad \text{or} \quad \begin{aligned} c_1 + c_2 &= x\\ c_2 &= y\\ 0 &= 0\end{aligned} \quad \text{to obtain}$$

$c_2 = y \Rightarrow c_1 = x - c_2 = x - y$.

Hence for any $\vec{x} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \in W$ we have $\vec{x} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = (x-y)\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + y\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$;

that is, every vector in W can be written as a linear combination of vectors in B. Hence B spans W and Span B = W.

Since B is a linearly independent set and spans W, it is a basis for W.

Q.E.D.

EXAMPLE (THEOREM) #2. B = $\{\hat{e}_1,\ldots,\hat{e}_n\}$ where $\hat{e}_i = \underbrace{[0,\ldots,0,1,0,\ldots,0]^T}_{\longleftarrow\ i^{th}\ \text{slot}\ \longrightarrow}$ is a basis of $\mathbf{R}^{n}$.

THEOREM #3. Let B = $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\} \subseteq W \subseteq V$ where W is a subspace of the vector space V. Then B is a basis of W iff $\forall\, \vec{x} \in W$, $\exists!\ c_1, c_2, \ldots, c_n$ such that $\vec{x} = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_k\vec{x}_k$.

The values of $c_1, c_2, \ldots, c_n$ associated with each $\vec{x}$ are called the coordinates of $\vec{x}$ with respect to the basis B = $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\}$. Given a basis, finding the coordinates of $\vec{x}$ for any given vector is an important problem..

Although a basis set is not unique, if there is a finite basis, then the number of vectors in a basis set isunique.

THEOREM #4. If B = $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ is a basis for a subspace W in a vector space V, then every basis set for W has exactly k vectors.

DEFINITION #2. The number of vectors in a basis set for a subspace W of a vector space V is the **dimension** of W. If the dimension of W is k, we write **dim W = k**.

THEOREM #5. The dimension of $\mathbf{R}^n$ over $\mathbf{R}$ (and the dimension of $\mathbf{C}^n$ over $\mathbf{C}$) is n.

Proof idea. Exhibit a basis and prove that it is a basis. (See Example (Theorem) #2)

EXERCISES on Basis Sets and Dimension

EXERCISE #1. True or False.

_____ 1. If B = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_n$ } $\subseteq$ V where is a vector space, then B is a basis of W if
B is linearly independent and B spans W (i.e. Span B = W)

_____ 2. To show that B is a spanning set using DUD we must show that an arbitrary vector
$\vec{b} \in$ W can be written as a linear combination of the vectors in B

_____ 3. To show that B = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_n$ } $\subseteq$ V where V is a vector space is a spanning set we
must show that for an arbitrary vector $\vec{b} \in$ W the vector equation
$c_1 \vec{x}_1 + c_2 \vec{x}_2 + \cdots + c_n \vec{x}_n = \vec{b}$ in the unknown variables $c_1, c_2, ..., c_n$ always has at least one solution.

_____ 4. If B = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_n$ } $\subseteq$ V where V is a vector space, then B is a basis of W iff
$\forall \vec{x} \in W, \exists! \; c_1, c_2, ..., c_n$ such that $\vec{x} = c_1\vec{x} + c_2\vec{x} + \cdots + c_n\vec{x}$.

_____ 5. B = { $[1,0,0]^T$, $[1,1,0]^T$ } is a basis for W = { $[x, y, 0]^T$: x,y $\in$ **R** }.

_____ 6. If B = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_n$ } $\subseteq$ W $\subseteq$ V where W is a subspace of the vector space V and B is a basis of W so that $\forall \vec{x} \in W, \exists! \; c_1, c_2, ..., c_n$ such that $\vec{x} = c_1\vec{x} + c_2\vec{x} + \cdots + c_n\vec{x}$, then the values of $c_1, c_2, ..., c_n$ associated with each $\vec{x}$ are called the coordinates of $\vec{x}$ with respect to the basis B.

_____ 7. A basis set for a vector space is not unique.

_____ 8. If B = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_n$ } is a basis for a subspace W in a vector space V, then every basis set for W has exactly n vectors.

_____ 9. The number of vectors in a basis set for a vector space V is called the dimension of V.

_____ 10. If the dimension of V is n, we write dim V = n.

_____ 11. The dimension of $\mathbf{R}^n$ over **R**.

_____ 12. The dimension of $\mathbf{C}^n$ over **C** is n.

EXERCISE #2. Show that B = { $[1,0,0]^T$, $[2,1,0]^T$ } is a basis for W = { $[x, y, 0]^T$: x,y $\in$ **R** }.

EXERCISE #3. Show that B = $\{\hat{e}_1, ..., \hat{e}_n\}$ where $\hat{e}_i = \underbrace{[0,...,0,1,0,...,0]^T}_{\longleftarrow i^{th} \text{ slot} \longrightarrow}$ is a basis of $\mathbf{R}^{n.}$.

EXERCISE #4. Show that the dimension of $\mathbf{R}^n$ over **R** is n.

EXERCISE #5. Show that the dimension of $\mathbf{C}^n$ over **C** is n.

LINEAR CLASS NOTES:
A COLLECTION OF HANDOUTS FOR
REVIEW AND PREVIEW
OF LINEAR THEORY
INCLUDING FUNDAMENTALS OF
LINEAR ALGEBRA

CHAPTER 6

Preview of the Theory of Abstract Linear Equations of the First Kind

In this handout, we review our preview of **linear operator** theory in the previous chapter. The most important examples of linear operators are differential and integral operators and operators defined by matrix multiplication. These arise in many applications. **Lumped parameter systems** (e.g., linear circuits and mass spring systems) have a finite number of state variables and give rise to **discrete operators** defined by matrices on finite dimensional vector spaces such as $\mathbf{R}^n$. Differential and integral equations (e.g., Maxwell's equations and the Navier-Stokes equation) are used to model **distributed (continuum) systems** having an infinite number of state variables and require infinite dimensional vector spaces (i.e., **function spaces**). That is, we have differential and integral operators on function spaces.

Even without covering any topics in differential equations, your background in calculus should be sufficient to understand discrete and continuous operators as linear operators on vector spaces.

A function or map T from one vector space V to another vector space W is often called an **operator**. If we wish to think geometrically (e.g., if V and W are $\mathbf{R}^2$ or $\mathbf{R}^3$) rather than algebraically we might call T a transformation.

DEFINITION 1. Let V and W be vector spaces over the same field **K**. An operator $T:V \to W$ is said to be **linear** if for all $\vec{x}, \vec{y} \in V$ and scalars α,β, it is true that

$$T(\alpha \vec{x} + \beta \vec{y}) = \alpha T(\vec{x}) + \beta T(\vec{y}). \tag{1}$$

THEOREM 1. Let V and W be vector spaces over the same field **K**. An operator $T: V \to W$ is linear if and only if the following two properties are true:

i) $\vec{x}, \vec{y} \in V$ implies $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y})$ (2)

ii) α a scalar and $\vec{x} \in V$ implies $T(\alpha \vec{x}) = \alpha T(\vec{x})$. (3)

EXAMPLE 1 Let the **operator** $T:\mathbf{R}^n \to \mathbf{R}^m$ be defined by **matrix multiplication** of the column vector $\vec{x}$ by the $m\times n$ matrix A; that is, let

$$T(\vec{x}) =_{df} \underset{m\times n}{A} \underset{n\times 1}{\vec{x}} \tag{4}$$

where $\underset{n\times1}{\vec{x}} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \in \mathbf{R}^{n\times1} = \mathbf{R}^n$ and $\underset{m\times n}{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdot & \cdot & \cdot & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdot & \cdot & \cdot & a_{2,n} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ a_{m,1} & a_{m,1} & \cdot & \cdot & \cdot & a_{m,n} \end{bmatrix} \in \mathbf{R}^{m\times n}$.

Then T is a linear operator.

EXAMPLE 2 Let I = (a,b). The operator $D:C^1(I,\mathbf{R}) \to C(I,\mathbf{R})$ be defined by

$$D(f) =_{df} \frac{df}{dx} \tag{5}$$

where $f \in C^1(I,R) = \{f:I \to \mathbf{R}: \frac{df}{dx}$ exists and is continuous on I$\}$ and $C(I,R) = \{f:I \to \mathbf{R} :$ f is continuous on I$\}$. Then D is a linear operator. We may restrict D to $\mathcal{A}(I, \mathbf{R}) = \{f:I \to \mathbf{R}:$f is analytic on I$\}$ so that $D:\mathcal{A}(I, \mathbf{R}) \to \mathcal{A}(I, \mathbf{R})$ maps a vector space back to itself.

DEFINITION #2. Let T:V→W be a mapping from a set V to a set W. The set $R(T) = \{ y \in W$: there exists an $x \in V$ such that $y = T(x) \}$ is called the **range** of T. If W has an additive structure (i.e., a binary operation which we call addition) with an additive identity which we call 0, then the set $N(T) = \{x \in V: T(x) = 0\}$ is called the **null set** of T (or nullity of T).

If T is a linear operator from a vector space V to another vector space W, we can say more.

THEOREM #2 Let T:V→W be a linear operator from a vector space V to a vector space W. The range of T $R(T) = \{ \vec{y} \in W$: there exists an $\vec{x} \in V$ such that $\vec{y} = T(\vec{x}) \}$, is a subspace of W and the null set of T, $N(T) = \{ \vec{x} \in V: T(\vec{x}) = 0\}$ is a subspace of V.

We rename these sets.

DEFINITION #3. Let T:V→W be a linear operator from a vector space V to another vector space W. The set $R(T) = \{ \vec{y} \in W$: there exists an $\vec{x} \in V$ such that $\vec{y} = T(\vec{x}) \}$ is called the **range space** of T and the set $N(T) = \{ \vec{x} \in V: T(\vec{x}) = 0\}$ is called the **null space** of T. .

Handout #2 | INTRODUCTION TO ABSTRACT LINEAR MAPPING PROBLEMS | Professor Moseley

We consider the (abstract) equation (<u>of</u> <u>the</u> <u>first</u> <u>kind</u>)

$$T(\vec{x}) = \vec{b} \qquad (\textbf{Nonhomogeneous}) \tag{1}$$

where T is a linear operator from the vector space V to the vector space W ($T:V \to W$). We view (1) as a mapping problem; that is, we wish to find those $\vec{x}$'s that are mapped by T to $\vec{b}$.

<u>THEOREM #1</u>. For the <u>nonhomogeneous</u> equation (1) there are three possibilities:
1) There are no solutions.
2) There is exactly one solution.
3) There are an infinite number of solutions.

<u>THEOREM #2</u>. For the <u>homogeneous</u> equation

$$T(\vec{x}) = \vec{0} \qquad (\textbf{Homogeneous}) \tag{2}$$

there are only two possibilities:
1) There is exactly one solution, namely $\vec{x} = \vec{0}$; that is the **<u>null space</u>** of T (i.e. the set of vectors that are mapped into the zero vector) is $N(T) = \{\vec{0}\}$.
2) There are an infinite number of solutions. If the null space of T is finite dimensional, say has dimension k ε **N**, then the **general solution** of (2) is of the form

$$\vec{x} = c_1\,\vec{x}_1 + \cdots + c_k\,\vec{x}_k = \sum_{i=1}^{k} c_i\,\vec{x}_i \tag{3}$$

where $B = \{\vec{x}_1, \cdots, \vec{x}_k\}$ is a basis for N(T) and c_i, i=1,...,k are arbitrary constants.

<u>THEOREM #3</u>. The nonhomogeneous equation (1) has at least one solution if $\vec{b}$ is contained in the **range space** of T, R(T), (the set of vectors $\vec{w} \in W$ for which there exist $\vec{v} \in V$ such that $T[\vec{v}] = \vec{w}$). If this is the case, then the **general solution** of (1) is of the form

$$\vec{x} = \vec{x}_p + \vec{x}_h \tag{4}$$

where $\vec{x}_p$ is a **<u>particular</u>** (i.e. any specific) **solution** to (1) and $\vec{x}_h$ is the **general** (e.g. a parametric formula for all) **solution(s)** of (2). If N(T) is finite dimensional then

$$\vec{x} = \vec{x}_p + \vec{x}_h = \vec{x}_p + c_1\,\vec{x}_1 + \cdots + c_k\,\vec{x}_k = \vec{x}_p + \sum_{i=1}^{k} c_i\,\vec{x}_i \tag{5}$$

where B = { $\vec{x}_1,\cdots,\ \vec{x}_k$ } is a basis of N(T). For the examples, we assume some previous knowledge of **determinants** and **differential equations**. Even without this knowledge, you should get a feel for the theory. And if you lack the knowledge, you may wish to reread this handout after obtaining it.

EXAMPLE 1 OPERATORS DEFINED BY MATRIX MULTIPLICATION

We now apply the general linear theory to operators defined by matrix multiplication. We look for the unknown column vector $\vec{x} = [x_1,x_2,\cdots,x_n]^T$. (We use the **transpose** notation on a **row vector** to indicate a **column vector** to save space and trees.) We consider the operator $T[\vec{x}] = A\vec{x}$ where A is an m×n matrix.

THEOREM 4. If $\vec{b}$ is in the column (range) space of the matrix (operator) A, then the **general solution** to the **nonhomogeneous system of algebraic equation(s)**

$$\underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{b}} \tag{6}$$

can be written in the form

$$\vec{x} = \vec{x}_p + c_1\vec{x}_1 + \cdots + c_k\vec{x}_k = \vec{x}_p + \sum_{i=1}^{k} c_i\vec{x}_i \tag{7}$$

where $\vec{x}_p$ is a **particular** (i.e. any) **solution** to (6) and

$$\vec{x}_h = c_1\vec{x}_1 + \cdots + c_k\vec{x}_k = \sum_{i=1}^{k} c_i\vec{x}_i \tag{8}$$

is the general solution (i.e. a parametric formula for all solutions) to the **complementary homogeneous equation**

$$\underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{0}} \tag{9}$$

Here B = { $\vec{x}_1,\cdots,\vec{x}_k$ } is a basis for the null space N(T) (also denoted by N(A)) which has dimension k. All of the vectors $\vec{x}_p,\ \vec{x}_1,\cdots,\vec{x}_k$ can be founded together using the computational technique of Gauss Elimination. If N(T) = { $\vec{0}$ }, then the unique solution of $\underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{b}}$ is $\vec{x}_p$ (and the unique solution to $\underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{0}}$ is $\vec{x}_h = \underset{n\times 1}{\vec{0}}$).

THEOREM 5. If n = m, then we consider two cases (instead of three) for equation (6):

1) det A ≠ 0 so that A is **nonsingular**; then the matrix A has a unique inverse, A^{-1} (which is almost never computed), and for any $\vec{b} \in \mathbf{R}^m$, $\underset{n\times n}{A}\ \underset{n\times 1}{\vec{x}} = \underset{n\times 1}{\vec{b}}$ always has the unique solution

$\vec{x} = A^{-1}\vec{b}$. Thus the operator $T(\vec{x}) = A\vec{x}$ is one-to-one and onto so that any vector $\vec{b}$ is always in the range space R(A) and only the vector $\vec{x} = A^{-1}\vec{b}$. maps to it. Again, the matrix A defines an operator that is a one-to-one and onto mapping from $\mathbf{R}^n$ to $\mathbf{R}^n$ (or $\mathbf{C}^n$ to $\mathbf{C}^n$).

2) det A = 0 so that A is **singular**; then either there is no solution or if there is a solution, then there are an infinite number of solutions. Whether there is no solution or an infinite numbers of solutions depends on $\vec{b}$, specifically, on whether $\vec{b}\ \varepsilon R(A)$ or not. The operator defined by tha matrix A is not one-to-one or onto and the dimension of N(A) is greater than or equal to one.

EXAMPLE 2 LINEAR DIFFERENTIAL EQUATIONS

To avoid using x as either the independent or dependent variable, we look for the unknown function u (dependent variable) as a function of t (independent variable). We let the domain of u be I = (a,b) and think of the function u as a vector in an (infinite dimensional) vector (function) space.

THEOREM 6. If g is in the range space R(L) of the linear differential operator L (i.e. $g\ \varepsilon\ R(L)$) then the **general solution to the nonhomogeneous equation**

$$L[\, u(t)\,] = g(t) \qquad \forall\, t \in I \tag{10}$$

can be written in the form

$$u(t) = u_p(t) + u_h(t) \tag{11}$$

where u_p is a **particular solution** to (10) and u_h is the **general solution to the homogeneous equation**

$$L[\, u(t)\,] = 0 \qquad \forall\, t \in I \tag{12}$$

Special cases:

1) $L[\, u(t)\,] = u'' + p(t)u' + q(t)u$. Second Order Scalar Equation.

For this case, we let I = (a,b) and $L: \mathcal{A}(I,R) \to \mathcal{A}(I,R)$. It is known that the dimension of the null space is two so that

$$u_h(t) = c_1 u_1(t) + c_2 u_2(t).$$

2) $L[\, u(t)\,] = p_o(t)\dfrac{d^n u}{dt^n} + \cdots + p_n(t)\, u(t)$ n^{th} Order Scalar Equation.

Again we let I = (a,b) and $L: \mathcal{A}(I,R) \to \mathcal{A}(I,R)$. For this case, the dimension of the null space is n so that

$$u_h(t) = c_1 u_1(t) + \cdots + c_n u_n(t) = \sum_{i=1}^{n} c_i\, u_i(t).$$

3) $L[\,\vec{u}(t)\,] = \dfrac{d\vec{u}}{dt} - \underset{n\mathrm{x}n}{P}(t)\vec{u}(t)$ First Order System ("Vector" Equation)

Again we let I = (a,b), but now $L:\mathcal{A}(I,\mathbf{R}^n) \to \mathcal{A}(I,\mathbf{R}^n)$ where $\mathcal{A}(I,\mathbf{R}^n) =$ $\{\vec{u}(t) : I \to R^n\}$;

that is the set of all time varying "vectors". Here the word "vector" means an n-tuple of functions. We replace (10) with

$$L[\,\vec{u}(t)\,] = \vec{g}(t)$$

and (12) with

$$L[\,\vec{u}(t)\,] = \vec{0}\,.$$

Then

$$\vec{u}(t) = \vec{u}_p(t) + \vec{u}_h(t)$$

where

$$\vec{u}_h(t) = c_1\, \vec{u}_1(t) + \cdots + c_n\, \vec{u}_n(t)$$ (i.e. the null space is n dimensional).

A SERIES OF CLASS NOTES TO INTRODUCE LINEAR AND NONLINEAR PROBLEMS
TO ENGINEERS, SCIENTISTS, AND APPLIED MATHEMATICIANS

LINEAR CLASS NOTES:
A COLLECTION OF HANDOUTS FOR
REVIEW AND PREVIEW
OF LINEAR THEORY
INCLUDING FUNDAMENTALS OF
LINEAR ALGEBRA

CHAPTER 7

Introduction to

Determinants

1. Introduction to Computation of Determinants

2. Computation Using Laplace Expansion

3. Computation Using Gauss Elimination

4. Introduction to Cramer's Rule

Rather than give a fairly complicated definition of the determinant in terms of minors and cofactors, we focus only on two methods for computing the determinant function det:$\mathbf{R}^{n\times n}\to\mathbf{R}$ (or det:$\mathbf{C}^{nxn}\to\mathbf{C}$). Let $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. The we define det(A) = ad−bc. Later we will show that $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} a & -b \\ -c & d \end{bmatrix} = \frac{1}{\det A}\begin{bmatrix} a & -b \\ -c & d \end{bmatrix}$. For $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$), we develop two methods for computing det(A): Laplace Expansion and Gauss Elimination

But first, we give (without proof) several properties of determinants that aid in their evaluation.

THEOREM. Let $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$). Then

1. (ERO's of type 1) If B is obtained from A by exchanging two rows, then det(B) = − det(A).
2. (ERO's of type 2) If B is obtained from A by multiplying a row of A by a≠0, then det(B) = a det(A).
3. (ERO's of type 3) If B is obtained from A by replacing a row of A by itself plus a scalar multiple of another row, then det(B) = det(A).
4. If U is the upper triangular matrix obtained from A by Gauss elimination (forward sweep) using only ERO's of type 3, then det(A) = det(U)
5. If $U\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$) is upper triangular, then det(U) is equal to the product of the diagonal elements.
6. If A has a row (or column) of zeros, then det(A) = 0..
7. If A has two rows (or columns) that are equal, then det(A) = 0..
8. If A has one row (column) that is a scalar multiple of another row (column), then det(A) = 0..
9. det(AB) =det(A) det(B)..
10. If det(A) ≠ 0, then $\det(A^{-1}) = 1/\det(A)$.
11. $\det(A^T) = \det(A)$.

EXERCISES on Introduction of Computation of Determinants

EXERCISE #1. True or False.

_____ 1. If $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then det(A) = ad−bc

_____ 2.If $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} a & -b \\ -c & d \end{bmatrix}$.

_____ 3.If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$), there are (at least) two methods for computing det(A).

_____ 4. If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$), Laplace Expansionis one method for computing det(A)

_____ 5. 3.If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$), use of Gauss Elimination is one method for computing det(A).

______ 6. If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$) and B is obtained from A by exchanging two rows, then det(B) = − det(A).

______ 7. If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$) and B is obtained from A by multiplying a row of A by $a\neq 0$, then det(B) = a det(A).

______ 8. If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$) and B is obtained from A by replacing a row of A by itself plus a scalar multiple of another row, then det(B) = det(A).

______ 9. If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$) and U is the upper triangular matrix obtained from A by Gauss elimination (forward sweep) using only ERO's of type 3, then det(A) = det(U)

______ 10. If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$) and $U\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$) is upper triangular, then det(U) is equal to the product of the diagonal elements.

______ 11. If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$) and A has a row of zeros, then det(A) = 0.

______ 12. If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$) and A has a column of zeros, then det(A) = 0.

______ 13. If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$) and A has two rows that are equal, then det(A) = 0.

______ 14. If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$) and A has two columns that are equal, then det(A) = 0.

______ 15. If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$) and A has one row that is a scalar multiple of another row, then det(A) = 0.

______ 16. If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$) and A has one column that is a scalar multiple of another column, then det(A) = 0.

______ 17. If $A,B\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$) then det(AB) =det(A) det(B)..

______ 18. If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$) and det(A) ≠ 0, then $\det(A^{-1}) = 1/\det(A)$.

______ 19. If $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$), thjen $\det(A^T) = \det(A)$.

<u>EXERCISE #2</u>. Compute det A where $A=\begin{bmatrix}0 & 5\\ 0 & -1\end{bmatrix}$

<u>EXERCISE #3</u>. Compute det A where $A=\begin{bmatrix}1 & 2\\ 3 & -1\end{bmatrix}$

<u>EXERCISE #4</u>. Compute det A where $A=\begin{bmatrix}1 & i\\ i & -1\end{bmatrix}$

<u>EXERCISE #5</u>. Compute det A where $A=\begin{bmatrix}1 & i\\ i & 0\end{bmatrix}$

<u>EXERCISE #6</u>. Compute A^{-1} if $A=\begin{bmatrix}2 & 5\\ 3 & 7\end{bmatrix}$

<u>EXERCISE #7</u>. Compute A^{-1} if $A=\begin{bmatrix}2 & 1\\ 5 & 3\end{bmatrix}$

<u>EXERCISE #8</u>. Compute A^{-1} if $A=\begin{bmatrix}1 & i\\ i & 2\end{bmatrix}$

<u>EXERCISE #9</u>. Compute A^{-1} if $A=\begin{bmatrix}1 & i\\ -i & 0\end{bmatrix}$

<u>EXERCISE #10</u>. Compute A^{-1} if $A=\begin{bmatrix}0 & 5\\ 0 & -1\end{bmatrix}$

EXERCISE #11. Compute A^{-1} if $A=\begin{bmatrix}1 & 2\\ 3 & -1\end{bmatrix}$

<u>EXERCISE #12</u>. Compute A^{-1} if $A=\begin{bmatrix}1 & i\\ i & -1\end{bmatrix}$

We give an example of how to compute a determinant using Laplace expansion..

EXAMPLE. Compute det(A) where $A=\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$ using Laplace expansion..

Solution: Expanding in terms of the first row we havc

$$\det(A) = \begin{vmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{vmatrix} = 2\begin{vmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{vmatrix} - (-1)\begin{vmatrix} -1 & -1 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & 2 \end{vmatrix}$$

$$+(0)\begin{vmatrix} -1 & 2 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & 2 \end{vmatrix} - (0)\begin{vmatrix} -1 & 2 & -1 \\ 0 & -1 & 2 \\ 0 & 0 & -1 \end{vmatrix}$$

so the last two 3x3's are zero. Hence expanding the first remaining 3x3 in terms of the first row and the second interms of the first column we have

$$\det(A) = 2\left[\begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} - (-1)\begin{vmatrix} -1 & -1 \\ 0 & 2 \end{vmatrix} + (0)\begin{vmatrix} -1 & 2 \\ 0 & -1 \end{vmatrix}\right] - (-1)\left[\begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} - (0)\begin{vmatrix} -1 & 0 \\ -1 & 2 \end{vmatrix} + (0)\begin{vmatrix} -1 & 0 \\ 2 & -1 \end{vmatrix}\right]$$

$$= 2[(4-1) + (-2)] + [4 - 1] = 2 + 3 = 5$$

EXERCISES on Computation Using Laplace Expansion

<u>EXERCISE #1</u>. Using Laplace expansion, compute det A where A= $\begin{bmatrix} 3 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$

<u>EXERCISE #2</u>. Using Laplace expansion, compute det A where A= $\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 3 \end{bmatrix}$

<u>EXERCISE #3</u>. Using Laplace expansion, compute det A where A= $\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & -1 & 2 \end{bmatrix}$

<u>EXERCISE #4</u>. Using Laplace expansion, compute det A where A= $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 1 & 4 \end{bmatrix}$

<u>EXERCISE #5</u>. Using Laplace expansion, compute det A where A= $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 2 & 1 & 4 \end{bmatrix}$

<u>EXERCISE #6</u>. Using Laplace expansion, compute det A where A= $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 1 & 4 \end{bmatrix}$

We give an example of how to compute a determinant using Gauss elimination.

EXAMPLE. Compute det(A) where $A=\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$ using Gauss elimination.

Recall

THEOREM. Let $A\in\mathbf{R}^{n\times n}$ (or $\mathbf{C}^{n\times n}$). Then

3. (ERO's of type 3) If B is obtained from A by replacing a row of A by itself plus a scalar multiple of another row, then det(B) = det(A).

4. If U is the upper triangular matrix obtained from A by Gauss elimination (forward sweep) using only ERO's of type 3, then det(A) = det(U).

$$R_2+(1/2)R_1\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} \Rightarrow R_3+(2/3)R_2\begin{bmatrix} 2 & -1 & 0 & 0 \\ 0 & 3/2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$$

$$\Rightarrow R_3+(3/4)R_2\begin{bmatrix} 2 & -1 & 0 & 0 \\ 0 & 3/2 & -1 & 0 \\ 0 & 0 & 4/3 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & -1 & 0 & 0 \\ 0 & 3/2 & -1 & 0 \\ 0 & 0 & 4/3 & -1 \\ 0 & 0 & 0 & 5/4 \end{bmatrix}$$

Since only ERO's of Type 3 were used, we have det(A) = det(U) = 2(3/2)(4/3)(5/4) = 5.

EXERCISES on Computation Using Gauss Elimination

EXERCISE #1. Using Gauss elimination, compute det A where $A=\begin{bmatrix} 3 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$

EXERCISE #2. Using Gauss elimination, compute det A where $A=\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 3 \end{bmatrix}$

EXERCISE #3. Using Gauss elimination, compute det A where $A=\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & -1 & 2 \end{bmatrix}$

Cramer's rule is a method of solving $A\vec{x} = \vec{b}$ when A is square and the determinant of A which we denote by D $\neq$ 0. The good news is that we have a formula. The bad news is that, computationally, it is not efficient for large matrices and hence is never used when n > 3. Let A be an nxn matrix and A = $[a_{ij}]$. Let $\vec{x} = [x_1, x_2, ..., x_n]^T$ and $\vec{b}$ be nx1 column vectors and $\vec{x} = [x_i]$ and $\vec{b} = [b_i]$. Now let A_i be the matrix obtained from A by replacing the ith column of a by the column vector $\vec{b}$. Denote the determinant of A_i by D_i. Then $x_i = D_i/D$ so that $\vec{x} = [x_1, x_2, ..., x_n]^T = [D_i / D]^T$.

EXERCISES on Introduction to Cramer's Rule

EXERCISE #1. Use Cramer's rule to solve $A\vec{x} = \vec{b}$ where $A = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

EXERCISE #2. Use Cramer's rule to solve $A\vec{x} = \vec{b}$ where $A = \begin{bmatrix} 2i & -3 \\ 3 & 5i \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

EXERCISE #3.Use Cramer's rule to solve $A\vec{x} = \vec{b}$ where $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -3 \\ 2 & 1 & -1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix}$

EXERCISE #4. Use Cramer's rule to solve

$x_1 + x_2 + x_3 + x_4 = 1$

$x_1 + 2x_2 + x_3 = 0$

$x_3 + x_4 = 1$

$x_2 + 2x_3 + x_4 = 1$

A SERIES OF CLASS NOTES TO INTRODUCE LINEAR AND NONLINEAR PROBLEMS
TO ENGINEERS, SCIENTISTS, AND APPLIED MATHEMATICIANS

LINEAR CLASS NOTES:
A COLLECTION OF HANDOUTS FOR
REVIEW AND PREVIEW
OF LINEAR THEORY
INCLUDING FUNDAMENTALS OF
LINEAR ALGEBRA

CHAPTER 8

Inner Product and

Normed Linear Spaces

Since solving linear algebraic equations for a field **K** require only a finite number of exact algebraic steps, any field will do. However, actually carrying out the process usually involves approximate arithmetic and hence approximate solutions. In an infinite dimensional vector space the solution process often requires an infinite process that converges. Hence the vector space (or its field) must have additional properties. Physicists and engineers think of vectors as quantities which have **length** and **direction**. Although the notion of direction can be discusssed, the definition of a vector space does not include the concept of the length of a vector. The abstraction of the concept of length of a vector is called a **norm** and vector spaces (also called **linear spaces**) which have a norm (or length) are called **normed linear spaces**. Having the notion of length in a vector space gives us the notion of a **unit vector** (i.e. a vector of length one).

DEFINITION #1. A **normed linear space** is a real or complex vector space V on which a norm has been defined. A norm (or length) is a function $\|\cdot\|: V \to \mathbf{R}^+ = \{\alpha \in \mathbf{R} : \alpha \geq 0\}$ such that

1) $\|\vec{x}\| \geq 0, \quad \forall \vec{x} \in V$

   $\|\vec{x}\| = 0$ if and only if $\vec{x} = \vec{0}$

2) $\|\alpha\vec{x}\| = |\alpha|\|\vec{x}\| \quad \forall \vec{x} \in V \quad \forall$ scalars α

3) $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\| \quad \forall \vec{x}_1, \vec{x}_2 \in V$ (this is called the triangle in equality)

Note that the zero vector $\vec{0}$ is the only vector with zero length. For all other vectors $\vec{x}$ we have $\|\vec{x}\| > 0$. Hence for each non zero $\vec{x}$ we can define the unit vector

$$\vec{u} = \vec{u}(\vec{x}) = \frac{1}{\|\vec{x}\|}\vec{x} = \frac{\vec{x}}{\|\vec{x}\|}. \tag{1}$$

LEMMA #1. If $\vec{x} \in V$, $\|\vec{x}\| \neq 0$ $\in$V, and $\vec{u} = \vec{u}(\vec{x}) = \frac{1}{\|\vec{x}\|}\vec{x} = \frac{\vec{x}}{\|\vec{x}\|}$, then $\|\vec{u}\| = 1$.

Proof. Let $\vec{u} = \vec{u}(\vec{x}) = \frac{1}{\|\vec{x}\|}\vec{x} = \frac{\vec{x}}{\|\vec{x}\|}$. Then

Statement	Reason
$\|\vec{u}\| = \left\|\frac{1}{\|\vec{x}\|}\vec{x}\right\|$	Definition of $\hat{u}$
$= \left|\frac{1}{\|\vec{x}\|}\right|\|\vec{x}\|$	Property (2) above
$= \frac{\|\vec{x}\|}{\|\vec{x}\|}$	Algebraic Properties of **R**
$= 1$	Algebraic Properties of **R**.

THEOREM #1. If $\vec{x} \neq \vec{0}$, then $\vec{x}$ can be written as $\vec{x} = \|\vec{x}\|\vec{u}$ where $\vec{u}$ is a unit vector in the direction of $\vec{x}$ and $\|\vec{x}\|$ gives the length of $\vec{x}$.

Proof idea. That $\|\vec{x}\|$ is the length or norm of $\vec{x}$ follows from the definition of $\|\vec{x}\|$ as a norm or length. Since $\vec{x} \neq \vec{0}$, $\|\vec{x}\| \neq 0$ and we define $\vec{u}$ as before as the unit vector $\vec{u} = \vec{u}(\vec{x}) = \frac{1}{\|\vec{x}\|}\vec{x} = \frac{\vec{x}}{\|\vec{x}\|}$ we see that $\vec{u}$ is a positive scalar multiple of $\vec{x}$ so that it is pointed in the same "direction" as $\vec{x}$. To show that $\vec{x} = \|\vec{x}\|\vec{u}$ is left as an exercise. QED

The abstraction of the notion of **distance** between two points in a set (or vector space) is called a **metric**.

DEFINITION #2. A **metric space** is a set S on which a metric has been defined. A **metric** (or distance between) on S is a function $\rho{:}S\times S\to \mathbf{R}^{+}=\{\alpha\in\mathbf{R}\ \alpha\geq 0\}$ such that

1) $\rho(x,y)\geq 0\ \forall$f $x,y\in S$.
   $\rho(x,y)=0$ if and only if $x=y$.
2) $\rho(x,y) = \rho(y,x)$.
3) $\rho(x,z) \leq \rho(x,y) + \rho(y,z)$ (this is also called the triangle in equality)

THEOREM #2. Every normed vector space is a metric space with the metric $\rho(\vec{x},\vec{y}) = \|\vec{x}-\vec{y}\|$.

However, we note that a metric space need not be a vector space. Geometrically in $\mathbf{R}^3$, ρ is the distance between the tips of the position vectors $\vec{x}$ and $\vec{y}$. A metric yields the notion of a **topology**, but we need not develop this more general concept. However, to discuss approximate solutions, we do need the notion of **completeness**. Although it could be developed in a more general context, we are content to discuss **complete vector spaces**.

DEFINITION #3. A **Cauchy sequence of vectors** is a sequence $\{\vec{x}_n\}_{n=1}^{\infty}$ in a normed linear space V such that for any $\varepsilon>0$, there exists N with $||\,\vec{x}_n - \vec{x}_m\,|| < \varepsilon$ whenever $m,n>N$.

Thus $\vec{x}_m$ and $\vec{x}_n$ get close together when n and m are large.

DEFINITION #4. A sequence of vectors $\{\vec{x}_n\}_{n=1}^{\infty}$ in a vector space V is convergent if there exist $\vec{x}$ such that for any $\varepsilon>0$, there exists N with $||\,\vec{x}_n - \vec{x}\,|| < \varepsilon$ whenever $n>N$.

DEFINITION #5. A vector space is **complete** if every Cauchy sequence of vectors converges. A **Banach Space** is a complete normed linear space.

The concepts of metric (or topology) and completeness are essential for computing limits; for example, in the process of computing approximate solutions and obtaining error estimates. Completeness does for a vector space what **R** does for **Q** (which are just special cases). It makes sure that there are no holes in the space so that Cauchy sequences (that look like they ought to converge) indeed have a vector to converge to. If we wish to solve problems in a metric space S and S is not complete, we can construct the completion of S which we usually denote by $\overline{S}$. Then, since $\overline{S}$ is complete, we can obtain approximate solutions.

Recall that to determine if two vectors in $\mathbf{R}^3$ are perpendicular, we compute the **dot product**. The abstraction of the notion of dot product in an abstract vector space is called an **inner product**. Vector spaces on which an inner product is defined are called **inner product spaces**. As we will see, in an inner product space we have not only the notion of two vectors being **perpendicular** but also the notions of **length of a vector** and a new way to determine if a set of vectors is **linearly independent**.

DEFINITION #1. An inner product space is a real or complex vector space V on which an inner product is defined. A inner product is a function $(\cdot,\cdot)$: $V \times V \to \mathbf{R}$ such that

1) $(\vec{x},\vec{y}) = \overline{(\vec{y},\vec{x})} \quad \forall \vec{x},\vec{y} \in V$

   (the bar over the inner product indicates complex conjugate. If V is a real vector space, it is not necessary and we see that the inner product is commutative for real vector spaces.)

2) $(\alpha\vec{x},\vec{y}) = \alpha(\vec{x},\vec{y}) \quad \forall \vec{x},\vec{y} \in V$ and scalars α.

3) $(\vec{x}_1+\vec{x}_2,\vec{y}) = (\vec{x}_1,\vec{y}) + (\vec{x}_2,\vec{y}) \quad \forall \vec{x}_1,\vec{x}_2,\vec{y} \in V$.

   (Properties 2 and 3) say that the inner product is linear in the first slot.)

4) $(\vec{x},\vec{x}) \geq 0 \quad \forall \vec{x} \in V$

   $(\vec{x},\vec{x}) = 0 \quad \text{iff } \vec{x} = \vec{0}$

If V is an inner product space we define a mapping, $\|\cdot\| : V \to \mathbf{R}^+ = \{\alpha \in \mathbf{R} : \alpha \geq 0\}$ by

$$\|\vec{x}\| = \sqrt{(\vec{x},\vec{x})}. \tag{1}$$

THEOREM #1. $\|\vec{x}\| = \sqrt{(\vec{x},\vec{x})}$ given by (1) defines a norm on any inner product space and hence makes it into a normed linear space. (See the previous handout)

DEFINITION #2. A Hilbert space is a complete, inner product space.

Again, the concept of completeness in a vector space is an abstraction of what $\mathbf{R}$ does for $\mathbf{Q}$. $\mathbf{R}$ is complete; $\mathbf{Q}$ is not.

EXAMPLE (THEOREM). Let $V = \mathbf{R}^n$ and define the inner product by $(\vec{x},\vec{y}) = \vec{x}^T\vec{y} = \sum_{i=1}^{n} x_i y_i$ where $\vec{x}^T = [x_1,\ldots,x_n]^T$. Note that we can define the inner product in $\mathbf{R}^n$ in terms of matrix multiplication. Note also $(\vec{x},\vec{y}) = \vec{y}^T\vec{x}$. We can then prove (i.e. verify) that $(\vec{x},\vec{y}) = \vec{x}^T\vec{y}$ defines an inner product (i.e. satisfies the properties in the definition of an inner product).

Proof of 2). Let $\vec{x}^T = [x_1,\ldots,x_n]^T$, $\vec{y}^T = [y_1,\ldots,y_n]^T$ then

Statement	Reason
$(\alpha\vec{x},\vec{y}) = \left(\alpha\vec{x}\right)^T \vec{y}$	definition of inner product for $\mathbf{R}^n$.
$= (\alpha[x_1,\ldots,x_n])\begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}$	notation and definition of transpose
$= [\alpha x_1,\ldots,\alpha x_n]\begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}$	definition of scalar multiplication
$= (\alpha x_1)y_1 + (\alpha x_2)y_2 +\cdots+ (\alpha x_n)y_n$	definition matrix multiplication
$= \alpha(x_1y_1 + x_2y_2 +\cdots+ x_ny_n)$	properties real numbers
$= \alpha(\vec{x}^T\vec{y})$	definition of matrix multiplication
$= \alpha(\vec{x},\vec{y})$	definition of inner product

QED

Note that in $\mathbf{R}^n$ we have

$$\|\vec{x}\| = (\vec{x},\vec{x}) = \vec{x}^T\vec{x} = x_1^{\,2} + \cdots + x_n^{\,2} = \sum_{i=1}^{n} x_i^{\,2} \ . \tag{2}$$

In $\mathbf{R}^3$ we know that two vectors are **perpendicular** if their dot product is zero. We abstract this idea by defining two vectors in an inner product space to be **orthogonal** (rather than use the word perpendicular) if their inner product is zero.

DEFINITION #3. The row vectors $\vec{x} = [x_1,x_2,\ldots,x_n]$ and $\vec{y} = [y_1,y_2,\ldots,y_n]$ in $\mathbf{R}^n$ are **orthogonal** if (and only if) $(\vec{x},\vec{y}) = \vec{x}^T\vec{y} = 0$.

Pythagoras Extended. In $\mathbf{R}^3$ (or any real inner product space) we might define two vectors to be perpendicular if they satisfy the Pythagorean theorem; that is if

$$\|x\|^2 + \|y\|^2 = \|x - y\|^2 \,(= \rho(\vec{x},\vec{y}))\ . \tag{3}$$

Since $\vec{x} = \vec{y} + (\vec{x} - \vec{y})$, (3) may be rewritten as

$$
\begin{aligned}
(\vec{x},\vec{x}) + (\vec{y},\vec{y}) &= (\vec{x}-\vec{y},\vec{x}-\vec{y}) \\
&= (\vec{x},\vec{x}) - 2(\vec{x},\vec{y}) + (\vec{y},\vec{y}) \qquad (4)
\end{aligned}
$$

(since the inner product is commutative; $\mathbf{C}^n$ is different.)

THEOREM #2. $\vec{x}$ and $\vec{y}$ in $\mathbf{R}^n$ are perpendicular iff they are orthogonal.

DEFINITION #4. $\hat{u} = \dfrac{\vec{x}}{\|\vec{x}\|}$ is a **unit vector** (i.e. a vector of length one) in the direction of the nonzero vector $\vec{x}$ ($\vec{0}$ has no direction). Hence any nonzero vector $\vec{x}$ can be written as $\vec{x} = \|\vec{x}\| \dfrac{\vec{x}}{\|\vec{x}\|}$ where $\|\vec{x}\|$ is the magnitude or length and $\hat{u} = \dfrac{\vec{x}}{\|\vec{x}\|}$ is a unit vector in the direction of $\vec{x}$.

DEFINITION #5. The cosine of the **acute angle** θ ($0 \le \theta \le \pi$) between two nonzero vectors $\vec{x}$ and $\vec{y} \in \mathbf{R}^n$ is

$$
\cos(\theta) = \frac{(\vec{x},\vec{y})}{\|\vec{x}\|\|\vec{y}\|} = \frac{\vec{x}^T\vec{y}}{\|\vec{x}\|\|\vec{y}\|} \qquad (5)
$$

Note: This is often used as the geometric definition of dot product in $\mathbf{R}^3$.
To show that (5) does yield θ in $\mathbf{R}^3$ we first extend the concept of **projection**.

DEFINITION #6. The **vector projection** of a vector $\vec{b}$ in the direction of a non-zero vector $\vec{a}$ is given by

$$
p = \frac{(\vec{a},\vec{b})}{\|\vec{a}\|} = \|\vec{b}\|\cos\theta
$$

scalar

$$
\vec{p} = \left(\|\vec{b}\|\cos\theta\right)\left(\frac{\vec{a}}{\|\vec{a}\|}\right)
$$

magnitude of $\vec{p}$ unit vector giving direction of $\vec{p}$

$$
= \frac{(\vec{a},\vec{b})}{\|\vec{a}\|}\frac{\vec{a}}{\|\vec{a}\|} = \frac{(\vec{a},\vec{b})}{\|\vec{a}\|^2}\vec{a}
$$

The magnitude of $\vec{p}$ is called the **scalar projection** of $\vec{b}$ in the direction of $\vec{a}$

We first review the definition, a theorem, and the test for a subspace.

DEFINITION #1. Let W be a nonempty subset of a vector space V. If for any vectors $\vec{x}$, $\vec{y} \in$ W and scalars $\alpha,\beta \in \mathbf{K}$ (recall that normally the set of scalars **K** is either **R** or **C**), we have that $\alpha\,\vec{x} + \beta\,\vec{y} \in$ W, then W is a **subspace** of V.

THEOREM #1. A nonempty subset W of a vector space V is a subspace of V if and only if for $\vec{x}$, $\vec{y} \in$ V and $\alpha \in \mathbf{K}$ (i.e. α is a scalar) we have.

i) $\vec{x}$, $\vec{y} \in$ W implies $\vec{x} + \vec{y} \in$ W, and

ii) $\vec{x} \in$ W implies $\alpha\,\vec{x} \in$ W.

TEST FOR A SUBSPACE. Theorem #1 gives a good test to determine if a given subset of a vector space is a subspace since we can test the closure properties separately. Thus if $W \subset V$ where V is a vector space, to determine if W is a subspace,we check the following three points.

1) Check to be sure that W is nonempty. (We usually look for the zero vector since if there is $\vec{x} \in$W, then $0\,\vec{x} = \vec{0}$ must be in W. Every vector space and every subspace must contain the zero vector.)
2) Let $\vec{x}$ and $\vec{y}$ be arbitrary elements of W and check to see if $\vec{x} + \vec{y}$ is in W. (Closure of vector addition)
3) Let $\vec{x}$ be an arbitrary element in W and check to see if $\alpha\,\vec{x}$ is in W. (Closure of scalar multiplication).

DEFINITION #2 Let W_1 and W_2 be subspaces of a vector space V. Then the **sum of W_1 and W_2** is defined as $W_1 + W_2 = \{\vec{x}_1 + \vec{x}_2 : \vec{x}_1 \in W_1 \text{ and } \vec{x}_2 \in W_2\}$.

THEOREM #2 Let W_1 and W_2 be subspaces of a vector space V. Then the sum of W_1 and W_2 is subspace of V.

DEFINITION #3 Let W_1 and W_2 be subspaces of a vector space V such that $W_1 \cap W_2 = \{\vec{0}\}$. Then the sum of W_1 and W_2 defined as $W_1 + W_2 = \{\vec{x}_1 + \vec{x}_2 : \vec{x}_1 \in W_1 \text{ and } \vec{x}_2 \in W_2\}$ is a **direct sum** which we denote by $W_1 \oplus W_2$

THEOREM #3 Let W_1 and W_2 be subspaces of a vector space V and $V = W_1 \oplus W_2 = \{\vec{x}_1 + \vec{x}_2 : \vec{x}_1 \in W_1 \text{ and } \vec{x}_2 \in W_2\}$. Then for every vector $\vec{x}$ in V there exist unique vectors $\vec{x}_1 \in W_1$ and $\vec{x}_2 \in W_2$ such that $\vec{x} = \vec{x}_1 + \vec{x}_2$.

DEFINITION #4 Let W_1 and W_2 be subspaces of a real inner product space V with inner product $(\cdot,\cdot)$. W_1 and W_2 are said to be **orthogonal** to each other if $\forall$ + $\vec{y}$ x $\in W_1$ and $\forall$ x + $\vec{y}$ $\in W_2$, we have $(\vec{x}, \vec{y}) = 0$. We write $W_1 {}_2 W_2$.

THEOREM #4. If $A \in \mathbf{R}^{m\times n}$, then its row space is orthogonal to its null space and its column space is orthogonal to its left null space. We write $R(A^T)\ {}_2\ N(A)$ and $R(A)\ {}_2\ N(A^T)$.

DEFINITION #5 Let W be a subspace of a real inner product space V with inner product $(\cdot,\cdot)$. The **orthogonal complement** of W is the set $W^2 = \{ \vec{y} \in V: (\vec{x}, \vec{y}) = 0\ \forall\ \vec{x} \in W\}$.

THEOREM #5. Let W be a subspace of a real inner product space V with inner product $(\cdot,\cdot)$. Then the orthogonal complement of W, $W^2 = \{ \vec{y} \in V: (\vec{x}, \vec{y}) = 0\ \forall\ \vec{x} \in W\}$, is a subspace.

THEOREM #6. Let W be a subspace of a real inner product space V with inner product $(\cdot,\cdot)$. Then the orthogonal complement of W^2 is W. We write $(W^2)^2 = W$.

THEOREM #7. If $A \in \mathbf{R}^{m\times n}$, then its row space is the orthogonal complement of its null space and the null is the orthogonal complement to the row space. We write $R(A^T)^2 = N(A)$ and $N(A)^2 = R(A^T)$. Similarly, $R(A)^2 = N(A^T)$ and $N(A^T)^2 = R(A)$.

THEOREM #8. Let W be a nonempty subset of a real inner product space V with inner product $(\cdot,\cdot)$. Then V is the direct sum of W and W^2, $V = W \oplus W^2$.

THEOREM #9. If $A \in \mathbf{R}^{m\times n}$, then $\mathbf{R}^n = R(A^T) \oplus N(A)$ and $\mathbf{R}^m = R(A) \oplus N(A^T)$.

In vector spaces where the concept of length or **norm** (as well as direction) is available, we can talk about approximate solutions to the mapping problem

$$T(\vec{u}) = \vec{b} \tag{1}$$

where T is a linear operator from V to W; T:V→W where V and W are normed linear spaces. Let $\vec{u}_a \in V$ be an **approximate solution** of (1) and $\vec{u}_e \in V$ be the **exact solution** which we assume to exist and be unique. A measure of how good a solution $\vec{u}_a$ is is given by the **norm** or length of the **error vector** in V,

$$\vec{E}_v = \vec{u}_e - \vec{u}_a\,;$$

that is,

$$E_V = \|\vec{E}_v\| = \|\vec{u}_e - \vec{u}_a\|\,.$$

If T is invertible (e.g., if T:$\mathbf{R}^n \to \mathbf{R}^n$, is defined by a matrix, T($\vec{x}$) = A $\vec{x}$ and detA$\neq$0), then

$$E_V = \|\vec{E}_v\| = \|\vec{u}_e - \vec{u}_a\| = \|T^{-1}(\vec{b}) - T^{-1}(\vec{b}_a)\| = \|T^{-1}(\vec{b} - \vec{b}_a)\|$$

where $\vec{b}_a = T(\vec{u}_a)$. (The inverse of a linear operator, if it exists, is a linear operator.) By a well-known theorem in analysis,

$$\|T^{-1}(\vec{b} - \vec{b}_a)\| \leq \|T^{-1}\|\,\|\vec{b} - \vec{b}_a\|$$

where $||T^{-1}||$ is the **norm of the operator** T^{-1} which we assume to be finite. If an a priori "estimate" (i.e., bound) of $||T^{-1}||$, say $||T^{-1}|| \leq C$, can be obtained, then an "estimate of" (i.e., bound for) E_V can be obtained by first computing $\|\vec{b} - \vec{b}_a\|$. Even without an estimate of (bound for) $||T^{-1}||$, we may use

$$E = E_W = \|\vec{b} - \vec{b}_a\| = \|\vec{b} - T(\vec{u}_a)\|. \quad ||b - b_a|| = ||b - T(u_a)||$$

where

$$\vec{E}_W = \vec{b} - \vec{b}_a = \vec{b} - T(\vec{u}_a)\,;$$

is a measure of the error for $\vec{u}_a$. After all, if $\vec{u}_a = \vec{u}_e$, then $T(\vec{u}_a) = \vec{b}$ so that $E_W = 0$. We call E_W the **error vector** in W. Note that this error vector (and hence E) can always be computed whereas $\vec{E}_V$ usually can not. (If $\vec{E}_V$ is known, then the exact solution $\vec{u}_e = \vec{u}_a + \vec{E}_V$ is known and there is no need for an approximate solution.) In fact, E can be computed independent of whether (1) has a unique solution or not. We refer to a solution that minimizes $E = E_W$ as a **least error solution**. If (1) has one or more solutions, then these are all least error solutions since they all give E = 0. On the other hand, if (1) has no solution, then choosing $\vec{u}$ to minimize E gives a “best possible” solution. Under certain conditions, it is unique.

DEFINITION #1. Let S = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_k$} ⊆ W ⊆ V where W is a subspace of the inner product space V. Then S is said to be (pairwise) orthogonal if for all i,j we have { $\vec{x}_i$, $\vec{x}_j$) = 0 for i≠j. If S is a basis of W, it is called an orthogonal basis. (This requires that S does not contain the zero vector.) An orthogonal basis is said o be othonormal if for alli, $\|\vec{x}_i\| \neq 0$

If B = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_k$} ⊆V is an orthogonal basis for the inner product space V, then the coordinates for $\vec{x} \in V$ are particularly easy to compute. Let

$$\vec{x} = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_k\vec{x}_k \in V \quad . \tag{1}$$

To find c_i, take the inner product of both sides with $\vec{x}_j$.

$$\begin{aligned}
(\vec{x}_j,\vec{x}) &= (\vec{x}_j, c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_k\vec{x}_k) \ .\\
(\vec{x}_j,\vec{x}) &= (\vec{x}_j, c_1\vec{x}_1) + (\vec{x}_j, c_2\vec{x}_2) + \cdots + (\vec{x}_j, c_k\vec{x}_k)\\
(\vec{x}_j,\vec{x}) &= c_1(\vec{x}_j,\vec{x}_1) + c_2(\vec{x}_j,\vec{x}_2) + \cdots + c_k(\vec{x}_j,\vec{x}_k)\\
(\vec{x}_j,\vec{x}) &= c_j(\vec{x}_j,\vec{x}_j)
\end{aligned} \tag{2}$$

so that

$$c_j = \frac{(\vec{x}_j,\vec{x})}{(\vec{x}_j,\vec{x}_j)} \ . \tag{3}$$

The concepts of a basis and orthogonal basis can be extended to infinite dimensional spaces. We first extend the concepts of linear independence and spanning sets.

DEFINITION #2. An infinite set S in a vector space V is linearly independent if every finite subset of S is linearly independent. Thus the countable set { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_k$,...} is linearly independent if Sn = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_k$} is linearly independent for all n∈**N**.

DEFINITION #3. Let the countable set S = { $\vec{x}_1$, $\vec{x}_2$,..., $\vec{x}_k$,...} ⊆ W ⊆ V where W is a subspace of a vector space V. S is a Hamel spanning set for W if for all $\vec{x} \in W$, there exists n∈**N** and c_1, c_2, ..., c_n such that $\vec{x} = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n$. If V is a topological vector space (e.g. a normed linear space), then S is a Schauder spanning set for W if for all $\vec{x} \in W$, there exist c_1, c_2, ..., c_n,... such that $\vec{x} = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n + \cdots = \sum_{n=1}^{\infty} c_n\vec{x}_n$.

DEFINITION #3. Let $B = \{\vec{x}_1, \vec{x}_2, ..., \vec{x}_n\} \subseteq W \subseteq V$ where W is a subspace of a vector space V. B is a Hamel basis for W if it is linearly independent and a Hamel spanning set for W. If V is a topological vector space (e.g. a normed linear space), then B is a Schauder basis for W if it is linearly independent and a Schauder spanning set for W..

EXAMPLE. 1) Let $B = \{1, x, x^2, x^3, ...\}$. Then B is a Hamel basis for the set of all polynomials P (R,R) and a Scauder basis for the set of all analytic functions with an infinite radius of convergence about x = 0. Note that both of these spaces are infinite dimensional.
2) Let $B = \{\vec{x}_1, \vec{x}_2, ..., \vec{x}_n, ...\}$ be an infinite linearly independent set in a real topological vector space. Then B is a Hamel basis for the subspace $W_1 = \{\vec{x} = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n : c_1, c_2, ,...c_n \in \mathbf{R}\}$ of V and B is a Schauder basis for the subspace $W_2 = \{\vec{x} = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n + : c_1, c_2, ,...c_n, ... \in \mathbf{R}$ where the series converges$\}$ of V. Again, note that both of these spaces are infinite dimensional.

If $B = \{\vec{x}_1, \vec{x}_2, ..., \vec{x}_n, ...\} \subseteq V$ is an orthogonal basis for the Hilbert space H, then the coordinates for $\vec{x} \in H$ are particularly easy to compute. Let

$$\vec{x} = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n + \cdots \in H \quad . \tag{4}$$

To find c_i, take the inner product of both sides with $\vec{x}_j$.

$$(\vec{x}_j, \vec{x}) = (\vec{x}_j, c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n + \cdots) \, .$$
$$(\vec{x}_j, \vec{x}) = (\vec{x}_j, c_1\vec{x}_1) + (\vec{x}_j, c_2\vec{x}_2) + \cdots + (\vec{x}_j, c_n\vec{x}_n) + \cdots$$
$$(\vec{x}_j, \vec{x}) = c_1(\vec{x}_j, \vec{x}_1) + c_2(\vec{x}_j, \vec{x}_2) + \cdots + c_n(\vec{x}_j, \vec{x}_n) + \cdots$$
$$(\vec{x}_j, \vec{x}) = c_j(\vec{x}_j, \vec{x}_j) \tag{5}$$

$$c_j = \frac{(\vec{x}_j, \vec{x})}{(\vec{x}_j, \vec{x}_j)} \, . \tag{6}$$

Note that this is the same formula as for the finite dimensional case.

A SERIES OF CLASS NOTES TO INTRODUCE LINEAR AND NONLINEAR PROBLEMS TO ENGINEERS, SCIENTISTS, AND APPLIED MATHEMATICIANS

LINEAR CLASS NOTES:
A COLLECTION OF HANDOUTS FOR
REVIEW AND PREVIEW
OF LINEAR THEORY
INCLUDING FUNDAMENTALS OF
LINEAR ALGEBRA

CHAPTER 9

More on

Matrix Inverses

1. Re-introduction to Matrix Inverses

2. Computation Using Gauss Elimination

4. Formula for a 2x2 Matrix

Recall that if A and B are square, then we can compute both AB and BA. Unfortunately, these may not be the same.

THEOREM #1. If n >1, then there exists A, B$\in\mathbf{R}^{n\times n}$ such that AB $\neq$ BA. Thus matrix multiplication is **not commutative**.

Thus AB=BA is not an identity. Can you give a **counter example** for n=2? (i.e. an example where AB $\neq$ BA.)

DEFINITION #1. For square matrices, there is a **multiplicative identity element**. We define the n×n matrix I by

$$\underset{mxn}{I} = \begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & \cdot & \cdot & \cdot & 0 \\ . & . & & & & . \\ . & . & & & & . \\ . & . & & & & . \\ 0 & 0 & \cdot & \cdot & \cdot & 1 \end{bmatrix}$$

One's down the diagonal. Zero's everywhere else.

THEOREM #2. We have $\underset{nxn}{A}\,\underset{nxn}{I} = \underset{nxn}{I}\,\underset{nxn}{A} = \underset{nxn}{A} \qquad \forall\, A \in \mathbf{K}^{nxn}$

DEFINITION #2. If there exists B such that AB = I., then B is a **right (multiplicative) inverse** of A. If there exists C such that CA = I., then C is a **left (multiplicative) inverse** of A. If AB = BA = I, then B is a (**multiplicative) inverse** of A and we say that A is **invertible**. If B is the only matrix with the property that AB = BA = I, then B is the **inverse** of A. If A has a unique inverse, then we say A is **nonsingular** and denote its inverse by A^{-1}.

THEOREM #3. Th identity matrix is its own inverse.

Later we show that if A has a right and a left inverse, then it has a unique inverse. Hence we prove that A is invertible if and only if it is nonsingular. Even later, we show that if A has a right (or left) inverse, then it has a unique inverse. Thus, even though matrix multiplication is not commutative, a right inverse is always a left inverse and is indeed the inverse. Some matrices have inverses; others do not. Unfortunately, it is usually not easy to look at a matrix and determine whether or not it has a (multiplicative) inverse.

THEOREM #4. There exist A, B$\in\mathbf{R}^{n\times n}$ such that A$\neq$I is invertible and B has no inverse.

INVERSE OPERATION. If B has a right and left inverse then it is a unique inverse ((i.e., $\exists B^{-1}$ such that $B^{-1}B = BB^{-1} = I$) and we can define **Right Division** AB^{-1} and **Left Division** $B^{-1}A$ of A by B (provided B^{-1} exists). But since matrix multiplication is not commutative, we do not know that these are the same. Hence $\frac{A}{B}$ is not well defined since no indication of whether we mean left or right division is given.

EXERCISES on Re-introduction to Matrix Inverses

EXERCISE #1. True or False.

_____ 1. If A and B are square, then we can compute both AB and BA.
_____ 2. If n >1, then there exists A, B$\in\mathbf{R}^{n\times n}$ such that AB $\neq$ BA.
_____ 3. Matrix multiplication is not commutative.
_____ 4. AB=BA is not an identity.
_____ 5. For square matrices, there is a multiplicative identity element, namely the n×n matrix I,

$$\text{given by } \underset{n\text{x}n}{I} = \begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & \cdot & \cdot & \cdot & 0 \\ . & . & & & & . \\ . & . & & & & . \\ . & . & & & & . \\ 0 & 0 & \cdot & \cdot & \cdot & 1 \end{bmatrix}.$$

_____ 6. $\forall A \in \mathbf{K}^{n\text{x}n}$ we have $\underset{n\text{x}n}{A}\ \underset{n\text{x}n}{I} = \underset{n\text{x}n}{I}\ \underset{n\text{x}n}{A} = \underset{n\text{x}n}{A}$
_____ 7. If there exists B such that AB = I., then B is a right (multiplicative) inverse of A.

_____ 8. If there exists C such that CA = I., then C is a left (multiplicative) inverse of A.
_____ 9. If AB = BA = I, then B is a multiplicative inverse of A and we say that A is invertible.
_____ 10. If B is the only matrix with the property that AB = BA = I, then B is the inverse of A.
_____ 11. If A has a unique inverse, then we say A is nonsingular and denote its inverse by A^{-1}.
_____ 12. The identity matrix is its own inverse.
_____ 13. If A has a right and a left inverse, then it has a unique inverse.
_____ 14. A is invertible if and only if it is nonsingular.
_____ 15. If A has a right (or left) inverse, then it has a unique inverse.
_____ 16. Even though matrix multiplication is not commutative, a right inverse is always a left inverse.
_____ 17. The inverse of a matrix is unique.
_____ 18. Some matrices have inverses; others do not.
_____ 19. It is usually not easy to look at a matrix and determine whether or not it has a (multiplicative) inverse.
_____ 20. There exist A, B$\in\mathbf{R}^{n\times n}$ such that A$\neq$I is invertible and B has no inverse

<u>EXERCISE #2</u>. Let $\alpha = 2$, $A = \begin{bmatrix} 1+i & 1-i \\ 1 & 0 \end{bmatrix}$, and $B = \begin{bmatrix} 1 & 0 \\ i & 1+i \end{bmatrix}$. Compute the following:

$\overline{A}$ = ____________. A^T =_____________. A^* = _____________. αA = _____________.

A+B =_____________. AB =______________.

<u>EXERCISE #3</u>. Let $\alpha = 3$, $A = \begin{bmatrix} i & 1-i \\ 0 & 1+i \end{bmatrix}$, and $B = \begin{bmatrix} 1 & 0 \\ i & 1+i \end{bmatrix}$. Compute the following:

$\overline{A}$ = ____________. A^T =_____________. A^* = _____________. αA = _____________.

A+B =_____________. AB =______________.

<u>EXERCISE #4</u>. Solve $\underset{2x2}{A}\ \underset{2x1}{\vec{x}} = \underset{2x1}{\vec{b}}$ where $A = \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} 1 \\ i \end{bmatrix}$.

<u>EXERCISE #5</u>. Solve $\underset{2x2}{A}\ \underset{2x1}{\vec{x}} = \underset{2x1}{\vec{b}}$ where $A = \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

EXERCISE #6 Solve $\underset{2x2}{A}\ \underset{2x1}{\vec{x}} = \underset{2x1}{\vec{b}}$ where $A = \begin{bmatrix} 1 & i \\ i & 0 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} 1 \\ i \end{bmatrix}$

We give an example of how to compute the inverse of a matrix A using Gauss-Jordan elimination (or Gauss-Jordan reduction). The procedure is to augment A with the identity matrix. Then use Gauss-Jordan to convert A into I. Magically, I is turned into A^{-1}. Obviously, if A is not invertible, this does not work.

<u>EXAMPLE</u>#1. Use Gauss-Jordan reduction to compute A^{-1} where $A=\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$.

$$\begin{array}{c} \\ R_2+(1/2)R_1 \\ \\ \\ \end{array}\left[\begin{array}{cccc|cccc} 2 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 2 & 0 & 0 & 0 & 1 \end{array}\right] \Rightarrow \begin{array}{c} \\ \\ R_3+(2/3)R_2 \\ \\ \end{array}\left[\begin{array}{cccc|cccc} 2 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 3/2 & -1 & 0 & 1/2 & 1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 2 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\Rightarrow \begin{array}{c} \\ \\ \\ R_4+(3/4)R_3 \end{array}\left[\begin{array}{cccc|cccc} 2 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 3/2 & -1 & 0 & 1/2 & 1 & 0 & 0 \\ 0 & 0 & 4/3 & -1 & 1/3 & 2/3 & 1 & 0 \\ 0 & 0 & -1 & 2 & 0 & 0 & 0 & 1 \end{array}\right] \Rightarrow \left[\begin{array}{cccc|cccc} 2 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 3/2 & -1 & 0 & 1/2 & 1 & 0 & 0 \\ 0 & 0 & 4/3 & -1 & 1/3 & 2/3 & 1 & 0 \\ 0 & 0 & 0 & 5/4 & 1/4 & 1/2 & 3/4 & 1 \end{array}\right]$$

We now divide by the pivots to make them all one.

$$\begin{array}{c} 1/2R_1 \\ 2/3R_2 \\ 3/4R_3 \\ 5/4R_4 \end{array}\left[\begin{array}{cccc|cccc} 2 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 3/2 & -1 & 0 & 1/2 & 1 & 0 & 0 \\ 0 & 0 & 4/3 & -1 & 1/3 & 2/3 & 1 & 0 \\ 0 & 0 & 0 & 5/4 & 1/4 & 1/2 & 3/4 & 1 \end{array}\right] \Rightarrow \left[\begin{array}{cccc|cccc} 1 & -1/2 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1 & -2/3 & 0 & 1/3 & 2/3 & 0 & 0 \\ 0 & 0 & 1 & -3/4 & 1/4 & 1/2 & 3/4 & 0 \\ 0 & 0 & 0 & 1 & 1/5 & 2/5 & 3/5 & 4/5 \end{array}\right]$$

We now make zeros above the pivots.

$$\begin{array}{c} \\ \\ R_3+3/4R_4 \\ \\ \end{array}\left[\begin{array}{cccc|cccc} 1 & -1/2 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1 & -2/3 & 0 & 1/3 & 2/3 & 0 & 0 \\ 0 & 0 & 1 & -3/4 & 1/4 & 1/2 & 3/4 & 0 \\ 0 & 0 & 0 & 1 & 1/5 & 2/5 & 3/5 & 4/5 \end{array}\right] \Rightarrow \begin{array}{c} \\ R2+2/3R_3 \\ \\ \\ \end{array}\left[\begin{array}{cccc|cccc} 1 & -1/2 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1 & -2/3 & 0 & 1/3 & 2/3 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2/5 & 4/5 & 6/5 & 3/5 \\ 0 & 0 & 0 & 1 & 1/5 & 2/5 & 3/5 & 4/5 \end{array}\right]$$

$$\Rightarrow \begin{array}{c} R_1+1/2R_2 \\ \\ \\ \\ \end{array}\left[\begin{array}{cccc|cccc} 1 & -1/2 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1 & -2/3 & 0 & 3/5 & 6/5 & 4/5 & 2/5 \\ 0 & 0 & 1 & 0 & 2/5 & 4/5 & 6/5 & 3/5 \\ 0 & 0 & 0 & 1 & 1/5 & 2/5 & 3/5 & 4/5 \end{array}\right] \Rightarrow \left[\begin{array}{cccc|cccc} 1 & -1/2 & 0 & 0 & 4/5 & 3/5 & 2/5 & 1/5 \\ 0 & 1 & -2/3 & 0 & 3/5 & 6/5 & 4/5 & 2/5 \\ 0 & 0 & 1 & 0 & 2/5 & 4/5 & 6/5 & 3/5 \\ 0 & 0 & 0 & 1 & 1/5 & 2/5 & 3/5 & 4/5 \end{array}\right]$$

Hence $A^{-1} = \begin{bmatrix} 4/5 & 3/5 & 2/5 & 1/5 \\ 3/5 & 6/5 & 4/5 & 2/5 \\ 2/5 & 4/5 & 6/5 & 3/5 \\ 1/5 & 2/5 & 3/5 & 4/5 \end{bmatrix}$. We will check.

$$AA^{-1} = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 4/5 & 3/5 & 2/5 & 1/5 \\ 3/5 & 6/5 & 4/5 & 2/5 \\ 2/5 & 4/5 & 6/5 & 3/5 \\ 1/5 & 2/5 & 3/5 & 4/5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Hence $A^{-1} = \begin{bmatrix} 4/5 & 3/5 & 2/5 & 1/5 \\ 3/5 & 6/5 & 4/5 & 2/5 \\ 2/5 & 4/5 & 6/5 & 3/5 \\ 1/5 & 2/5 & 3/5 & 4/5 \end{bmatrix}$ is indeed the inverse of A.

<u>EXAMPLE</u>#2. For a 2x2 we can do the computation in general. This means that we can obtain a formula for a 2x2. We can do this for a 3x3, but the result is not easy to remember and we are better off just using Gauss-Jordan for the particular matic of interest. Let $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. For the 2x2, we will assume that $a\neq 0$. We leave it to the exercises to show that the formula that we will derive for $a\neq 0$ also works for $a = 0$. Proceeding we first get zeros below the pivots.

$$R_2 - c/aR_1 \left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right] \Rightarrow \left[\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & d - \frac{c}{a}b & -\frac{c}{a} & 1 \end{array}\right] \text{ or } \left[\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & \frac{ad-cb}{a} & -\frac{c}{a} & 1 \end{array}\right]$$

Next we make the pivots all one. We now assume that $\det(A)= ad-bc \neq 0$ so that the matrix is nonsingular.

$$\begin{array}{c} 1/aR_1 \\ a/(ad-bc)R_2 \end{array} \left[\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & \frac{ad-cb}{a} & -\frac{c}{a} & 1 \end{array}\right] \Rightarrow R_1 - b/aR_2 \left[\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right]$$

$$\Rightarrow \left[\begin{array}{cc|cc} 1 & 0 & \frac{1}{a} - \frac{b}{a}\frac{c}{ad-bc} & -\frac{b}{a}\frac{a}{ad-bc} \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right] \Rightarrow \left[\begin{array}{cc|cc} 1 & 0 & \frac{ad-(bc-bc)}{a(ad-bc)} & -\frac{b}{ad-bc} \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right] \Rightarrow \left[\begin{array}{cc|cc} 1 & 0 & \frac{d}{a(ad-bc)} & -\frac{b}{ad-bc} \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right]$$

Hence $A^{-1} = \begin{bmatrix} \frac{d}{a(ad-bc)} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

EXERCISES on Computation Using Gauss-Jordan Elimination

<u>EXERCISE #1</u>. Using the formula $A^{-1} = \frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, compute A^{-1} when $A=\begin{bmatrix} 6 & 1 \\ 11 & 2 \end{bmatrix}$

EXERCISE #2. Using the formula $A^{-1} = \; = \dfrac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, compute A^{-1} where $A = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}$

EXERCISE #3. Using Gauss-Jordan elimination, compute A^{-1} where $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

EXERCISE #4. Using Gauss-Jordan elimination, compute A^{-1} where $A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix}$

EXERCISE #5. Let $A = \begin{bmatrix} 0 & b \\ c & d \end{bmatrix}$. Without using the formula $A^{-1} = \; = \dfrac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

use Gauss-Jordan to show that $A^{-1} = \; = \dfrac{1}{\det A}\begin{bmatrix} d & -b \\ -c & 0 \end{bmatrix}$. Thus you have proved that the formula

works even when a = 0.

EXERCISE #6. Using the formula $A^{-1} = \; = \dfrac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, compute A^{-1} where $A = \begin{bmatrix} 0 & 1 \\ 4 & 2 \end{bmatrix}$.

EXERCISE #7. Compute A^{-1} if $A = \begin{bmatrix} 2 & 5 \\ 3 & 7 \end{bmatrix}$

EXERCISE #8. Compute A^{-1} if $A = \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}$

EXERCISE #9. Compute A^{-1} if $A = \begin{bmatrix} 1 & i \\ i & 2 \end{bmatrix}$

EXERCISE #10. Compute A^{-1} if $A = \begin{bmatrix} 1 & i \\ -i & 0 \end{bmatrix}$

EXERCISE #11. Compute A^{-1} if $A = \begin{bmatrix} 0 & 5 \\ 0 & -1 \end{bmatrix}$

EXERCISE #12. Compute A^{-1} if $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$

EXERCISE #13. Compute A^{-1} if $A = \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$

EXERCISE #14. Using Gauss elimination, compute A^{-1} where $A = \begin{bmatrix} 3 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$

EXERCISE 15. Using Gauss elimination, compute A^{-1} where $A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & -1 & 2 \end{bmatrix}$

EXERCISE #16. Using Gauss elimination, compute A^{-1} where $A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

EXERCISE #17. Using Gauss elimination, compute A^{-1} where $A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$