Portfolio for Bachelor of Science in Mathematics
Indiana University East

Brett Hansen
Senior in Mathematics
Graduation: August 2015

April 27, 2015

Contents

1 Calculus I - III

We will begin our summaries of math courses with a brief summary of some of the topics covered in Calculus courses taken at the university level. Calculus is a fundamental tool with far reaching applications to the STEM fields and beyond.

Unless otherwise stated, information for the following summary of calculus is taken from [10].

1.1 Definition of Derivative

The *derivative* of a function $f(x)$ gives the instantaneous rate of change of the function, and the value of the derivative at x is the value of the slope of a line tangent to the graph of $f(x)$ at that point [9]. A formal definition from [5] is given below.

Definition 1.1. *For any given function f the* derivative *of f at the number x can be denoted $f'(x)$, and defined as follows:*

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

An alternative definition for the derivative of is given below.

$$f'(x) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}.$$

1.2 Rules for Differentiation

1.2.1 Derivatives of Poynomials and Exponential Functions

What follows are several rules for finding the derivative of common polynomial and exponential functions.

Theorem 1.2. *The derivative of a constant function c with respect to a variable x is as follows:*

$$\frac{d}{dx}(c) = 0.$$

Theorem 1.3. *The derivative of a x with respect to a variable x is as follows:*

$$\frac{d}{dx}(x) = 1.$$

This can be generalized for all multiples of x as:

$$\frac{d}{dx}(cx) = c.$$

The derivative of a constant function, and the derivative of the function x are both specialized cases of the power rule. The power rule allows us to differentiate any power of x and can be stated as follows:

Theorem 1.4. *The Power Rule:*

$$\frac{d}{dx}(cx^n) = cnx^{n-1}$$

From here, we must also define how to deal with the addition and subtraction operations relating to differentiation. The derivative of sums or differences is found by breaking the function down into individual chunks and finding the derivatives of their constituent parts.

Theorem 1.5. *Sum Rule:*

$$\frac{d}{dx}(f(x) + g(x)) = \frac{d}{dx}f(x) + \frac{d}{dx}g(x)$$

Likewise, there is very little difference between the derivative of a sum or of a difference. Difference Rule:

$$\frac{d}{dx}(f(x) - g(x)) = \frac{d}{dx}f(x) - \frac{d}{dx}g(x)$$

Next we must define a way to deal with exponentials and functions with a variable as an exponent.

Theorem 1.6. *Derivative of the exponential function:*

$$\frac{d}{dx}e^x = e^x$$

Please note that this is actually a specialized form of a more general derivative.

$$\frac{d}{dx}a^x = a^x ln(a).$$

In our first formula, ln(e) cancels out leaving only the first term

1.2.2 Product and Quotient rules

In order to differentiate the widest possible range of functions, there must be rules to deal with functions composed of products and quotients of other functions. Luckily a simple set of rules to do so exists.

Theorem 1.7. *Product Rule*

Let F(x) = f(x)g(x) where f(x) and g(x) are differentiable functions where

$$\frac{d}{dx}f(x) = f'(x)$$

and

$$\frac{d}{dx}g(x) = g'(x).$$

Then the derivative of F(x) is:

$$\frac{d}{dx}F(x) = f'(x)g(x) + f(x)g'(x)$$

After this we need a way to deal with Functions that are the quotient of other functions.

Theorem 1.8. *Quotient Rule*
Let F(x) = f(x)/g(x) where f(x) and g(x) are differentiable functions where

$$\frac{d}{dx}f(x) = f'(x)$$

and

$$\frac{d}{dx}g(x) = g'(x).$$

Then the derivative of F(x) is:

$$\frac{d}{dx}F(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}.$$

1.2.3 Chain Rule

Our basic toolbox of derivative tools is nearly complete. However, we need to have a tool to deal with function composition. This refers to functions of the form:

$$F(x) = f(g(x))$$

where f(x) is a differentiable function with the derivative f'(x) and g(x) is a differentiable function with the derivative g'(x).

Theorem 1.9. *Chain Rule:*

$$\frac{d}{dx}F(x) = \frac{d}{dx}f(g(x)) = f'(g(x))g'(x).$$

1.3 Definite integral

Integration at its most basic deal with finding the area under a curve. This has many applications because any continuous function can be thought of as tracing a curve. A definite integral is the area under a continuous curve with defined endpoints.

Theorem 1.10. *Definite Integral:*

$$\int_a^b f(x)dx = \lim_{n\to\infty}\sum_{i=0}^{n} f(x_i^*)\Delta x$$

where Δx is the size of the step between successive elements of f(x) and $f(x_i)$ denotes element i of f(x) between the limits of a and b.1

1.4 Techniques of integration

Integration has many rules which function similarly to the rules of differentiation.

This makes it possible for differentiation and integration to undo each other. This means that $\frac{d}{dx} \int f(x)dx = f(x)$

For non-trivial integrals, we have an array of techniques to enable us to find solutions. One of these techniques will be detailed in the sections to follow.

1.4.1 Parts

Integration by Parts is, in some ways a complement to the product rule from differentiation. It allows one to find an integral of a product. With continuous functions u and v it takes the form:

$$\int udv = uv - \int vdu.$$

This can be a little difficult to understand, so we can rewrite this with the example functions f(x) and g(x) to get:

$$\int f(x)g'(x)dx = f(x)g(x) - \int g(x)f'(x).$$

1.5 Partial Derivatives

Partial variables deal with functions of two or more variables. In the simplest case, it deals with a function of two variables.

Definition 1.11. *A function f of two variables is a rule that assigns to each ordered pair of real numbers x, y in a set D a unique real number denoted by f(x,y). The set D is the domain of f and its range is the set of values that f takes on, that is,* $f(x,y)|(x,y) \in D$.

Functions in 2 variables are graphed in the following way:

Definition 1.12. *If f is a function of two variables with domain D, then the graph of f is the set of all points (x, y, z) in* R^3 *such that z = f(x, y) and (x, y) is in D.*

In the same way that continuity is essential to differentiation in one variable, it is also essential to differentiation in multiple variables. We define continuity in the following way:

Definition 1.13. *A function f of two variables is called continuous at (a, b) if* $\lim_{(x,y)\to(a,b)} f(x,y) = f(a,b)$ *We say f is continuous on D if f is continuous at every point a, b in D .*

Now we have the essential definitions to define our partial derivatives.

Theorem 1.14. *Partial Derivative:*

If f is a function of two variables, its partial derivatives are the functions f_x and f_y defined by $f_x(x,y) = \lim_{h\to 0}, \dfrac{f(x+h,y)-f(x,y)}{h}$ and $f_y(x,y) = \lim_{h\to 0}, \dfrac{f(x,y+h)-f(x,y)}{h}$

2 Linear Algebra

Now we will move on to the field of Linear Algebra. Linear Algebra is at its most basic, a suite of tools, techniques and methodologies associated with large systems of linear equations.

Please note that unless otherwise stated, the information in the following summary of Linear Algebra is drawn from Schaum's Outline of Linear Algebra [7].

2.1 Systems of Linear Equations

The most basic topic of discussion in Linear Algebra is the idea of a System of Linear Equations. Linear equations are at their most simple, equations of the form $ax = b$ where a and b are constants, and x is a variable. This is a trivial problem to solve, but supposing we have more than one variable. Generally, our linear equations will take the form:

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b$$

Given this and this alone, it would be impossible to come upon a singular solution. However, in Linear Algebra, we see systems of these linear equations and use them together to come up with a solution. So for 3 variables, we might have something like this:

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 = b_2 \\ a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 = b_3 \end{aligned}$$

Theorem 2.1. *It is worth noting that a system of n equations in n unknowns is more likely to have a specific solution.*

If we were to have to write this whole thing out every single time we wanted to do something, it would get extremely tedious. As a result, we've developed a beautiful piece of shorthand: the matrix.

A matrix is an extremely powerful structure which contains the information of a system of equations. The matrix of our previous example looks like so:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$

This would be a simple *coefficient matrix* containing all the coefficients of the variables in our sample problem. There also exists another type of matrix commonly in use called the augmented matrix. This matrix is "augmented" by having the solution as a part of the matrix as well. It would look something like this

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & b_1 \\ a_{2,1} & a_{2,2} & a_{2,3} & b_2 \\ a_{3,1} & a_{3,2} & a_{3,3} & b_3 \end{bmatrix}$$

This gives us a very simple idea of what matrices are. In essence, they are just another, more concise way of demonstrating the relationships in a system of equations. Now that we have covered this, we can move on to why they're so powerful.

2.2 Echelon Forms

What makes matrices so useful is the fact that they can be easily and intuitively manipulated. We'll get into the exact details of the rules of matrix operations in a bit. But for now, all we need to know is that there exist certain prescribed rules that keep the system identical from start to finish.

Theorem 2.2. *Two matrices are identical if they have the same solution set*

To say that a matrix is in eschelon form is to say that it follows certain rules regarding its structure. There are 2 fundamental types of Eschelon form that we will be concerned with in this class: Row Eschelon form (REF) and Reduced Row Eschelon Form (RREF).

Firstly, Row Eschelon form is a matrix where all entries below the leading non-zero entries are zero. This can be a bit difficult to visualize, so here is how it looks in our example problem.

$$\begin{bmatrix} 1 & a_{1,2} & a_{1,3} & b_1 \\ 0 & 1 & a_{2,3} & b_2 \\ 0 & 0 & 1 & b_3 \end{bmatrix}$$

As you can see, this makes it easy to solve for the last variable (since these correspond to those original linear equations) and subsitute backwards to solve the system. However, there exists an identical matrix which has the solution even more plainly displayed.

The Reduced Row Eschelon Form of a matrix is where the leading entry of each row is a 1 and all other entries in the column are zero. In terms of our example, it would look like this:

$$\begin{bmatrix} 1 & 0 & 0 & b_1 \\ 0 & 1 & 0 & b_2 \\ 0 & 0 & 1 & b_3 \end{bmatrix}$$

If we take a look at this, we see quite plainly that b_3 contains the solution of x_3, and likewise that b_2 contains the solution of x_2 and lastly that b_1 contains the solution of x_1. As you can imagine, this simplicity gives us quite a lot of incentive to find the Reduced Row Eschelon Form of a Matrix. As a result, we have a fair number of methods to get us the Reduced Row Eschelon Form of a Matrix.

2.3 Matrix Operations and Inverse Matrices

Here we will cover the basic algebra of matrices.

2.3.1 Matrix Addition and Scalar Multiplication

We can add two matrices together, provided they have the same dimensions. It is done by adding up the corresponding coefficients in each matrix. For example, suppose we have a matrix A (our example matrix) and a matrix C, defined as

$$\begin{bmatrix} c_{1,1} & c_{1,2} & c_{1,3} & d_1 \\ c_{2,1} & c_{2,2} & c_{2,3} & d_2 \\ c_{3,1} & c_{3,2} & c_{3,3} & d_3 \end{bmatrix}$$

In this case, the result of $[A] + [C]$ would be :

$$\begin{bmatrix} a_{1,1} + c_{1,1} & a_{1,2} + c_{1,2} & a_{1,3} + c_{1,3} & b_1 + d_1 \\ a_{2,1} + c_{2,1} & a_{2,2} + c_{2,2} & a_{2,3} + c_{2,3} & b_2 + d_2 \\ a_{3,1} + c_{3,1} & a_{3,2} + c_{3,2} & a_{3,3} + c_{3,3} & b_3 + d_3 \end{bmatrix}$$

Please note that we can also simply multiply a matrix by a scalar c. In this case, we would simply multiply each entry in the matrix by our scalar. Note that multiplying a matrix by a scalar leaves you with an identical matrix of a different magnitude. For example, if we multiplied our example matrix A by c, we would get

$$\begin{bmatrix} ca_{1,1} & ca_{1,2} & ca_{1,3} & cb_1 \\ ca_{2,1} & ca_{2,2} & ca_{2,3} & cb_2 \\ ca_{3,1} & ca_{3,2} & ca_{3,3} & cb_3 \end{bmatrix}$$

2.3.2 Matrix Multiplication

We can also multiply two matrices together by multiplying the rows of one matrix by the columns of the second. In our example matrices A and C, AC would be

$$\begin{bmatrix} a_{11}c_{11} + a_{21}c_{12} + a_{31}c_{13} & a_{21}c_{12} + a_{22}c_{22} + a_{32}c_{23} & a_{31}c_{13} + a_{23}c_{32} + a_{33}c_{33} \\ . & . & . \\ . & . & . \end{bmatrix}$$

Obviously, this is a step where errors can creep in due to the sheer number of operations that need to be done. As such, it's important to double check your work when doing matrix multiplication.

It's interesting to note that many of the early uses of computers were to solve very large linear systems because doing lots of small calculations quickly is precisely where computers to excel.

2.3.3 Transpose of a Matrix

The transpose of a matrix A, denoted by $[A]^T$ is the matrix formed by swapping the rowspace and column space of the matrix. In our example matrix of

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$

The transpose would be:

$$\begin{bmatrix} a_{1,1} & a_{2,1} & a_{3,1} \\ a_{1,2} & a_{2,2} & a_{3,2} \\ a_{1,3} & a_{2,3} & a_{3,3} \end{bmatrix}$$

Note that the entries along the diagonal are not affected by transposition.

2.3.4 Inverse Matrix

An inverse matrix of a matrix A denoted by $[A]^{-1}$. A matrix's inverse "undoes it". That is to say, $(A)(A^{-1}) = I$ where I is the identity matrix. The identity matrix is the multiplicative identity of the matrix world. It is a matrix with 1's along the diagonal and 0's everywhere else. In a 3×3 case, the identity matrix would be:

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

Note that $AI = A$.

2.4 Determinants and Cramers Rule

In this section we'll be exploring determinants and a very nifty use for them.

2.4.1 Determinants

Every n by n (square) has a unique scalar called a determinant denoted by det(A). The determinant of a matrix gives us some clue about certain properties of the matrix such as diagonalizability or whether it is solvable.

In the trivial case of a 1×1 matrix, the determinant is just the single entry in the matrix.

In a 2×2 matrix of the form

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

has a determinant value $det(A) = ad - bc$.

For a 3×3 matrix A, defined as:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

has a determinant value $det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$.

For larger matrices, we find the determinant by breaking the matrix into smaller submatrices and finding the sum of those smaller determinants.

2.4.2 Cramer's Rule

Cramer's rule is a technique to find the solution of a linear system Ax = b. It is defined as follows:

Theorem 2.3. *For a system Ax = b, provided det(A) does not equal 0, the solution is given by*

$$x_1 = \frac{N_1}{det(A)}$$

$$x_2 = \frac{N_2}{det(A)}$$

$$...$$

$$x_n = \frac{N_n}{det(A)}$$

In this case, the "N" values are given as the determinant of that submatrix formed by the exclusion of that row and column.

3 Bridge to Abstract Mathematics

Notation has a huge impact on the math we can do. As we saw with matrices in our summary of Linear Algebra, proper notation gives us the structure required to think clearly about math and also give us the ability to cogently and simply pass along our arguments in such a way that an observer can easily follow our though processes. This summary of the "bridge to abstract mathematics" will be devoted to the notation, terminology and common techniques that appear in the course.

Please note that unless otherwise stated, the information in the following sections is gleaned from [3]. The examples within come from the same source as my own notes on the subject were lost in a move.

3.1 Truth Tables, Universal Statements, Existential Statements

We start our summary with the basic concept of the truth table. The truth table is a basic tool for logic which consists of two separate parts. First, there are columns representing the possible states of certain conditional statements, commonly denoted by letters such as P and Q which can be have the possible values of true or false. Then we have results columns which show the result of certain statements involving our P's and Q's.

Quickly we must do some definition to ensure that there is no confusion.

Definition 3.1. $\wedge$ *is the logical shorthand for and will be evaluated as true if the conditionals on both sides are evaluated as true.*

Definition 3.2. $\vee$ *is the logical shorthand for or, which will be evaluated as true if the conditionals on either side are evaluated as true.*

Please note that this is *not* the exclusive or, if both sides are true, the statement is evaluated as true. Last but not least, $\neg$ is the logical shorthand for not, which inverts the value of the statement made.

Here is an example of a Truth table for P or Q.

Example 3.3.

P	Q	$P \vee Q$
T	F	T
F	T	T
T	T	T
F	F	F

This concisely shows that the only way for the statement "P or Q" to be false is for P and Q to be false, with every other case making P or Q true.

Now let's take a look at a more complex truth table which makes use of the three logical operations we have defined.

Example 3.4.

P	Q	R	$(P \wedge Q) \vee \neg R$	$\neg(P \vee Q) \wedge R$
T	F	F	T	F
F	T	F	T	F
F	F	T	F	T
T	F	T	F	F
T	T	F	T	F
F	T	T	F	F
T	T	T	F	F
F	F	F	T	F

This shows the utility of truth tables because you can very easily see the truth value of a statement with only a glance at the table.

Let's take a brief look at two fundamental types of statements that we can make.

The universal statement is one which states something for a field of certain proscribed values. If you see something like $\forall X \in R$ which reads "For all X in the set of Real Numbers", you will know they are making a universal statement.

The existential statement is one which asserts the existence of a particular value that makes some statement true. If you see something like $\exists x \in Z$ which reads as "There exists some x in the set of Integers" then you will know you are dealing with an existential statement.

3.2 Sets and Set Operations

In the last section we briefly mentioned the idea of sets of numbers. It is of the utmost importance that we carefully define what constitutes a set.

Theorem 3.5. *A* set *is a well-defined collection of objects. A set can have between 0 and* inf *elements contained within it.*

Definition 3.6. *Any set which contains 0 elements is called the null set and is denoted by* $\emptyset$

Here is a list of many of the commonly used sets:
$\mathbb{Z}$: The set of all integers (whole numbers)
$\mathbb{Q}$: The set of all rational numbers (of the form a/b where $a, b \in \mathbb{Z}$)
$\mathbb{R}$: The set of all numbers which can be expressed in a decimal form
$\mathbb{C}$: The set of all numbers of the form $a + bi$ where $i = \sqrt{-1}$

Note that we can also break up these sets into other sets through modifiers. For example: $\mathbb{Z}^+$ is the set of all positive integers ($\forall z \in \mathbb{Z}^+, z > 0$)

There are several ways to define sets. The most fundamental way to define a finite set (one with a countable number of elements) is to simply list out its elements. For example, a finite set S could be defined as follows:

Example 3.7. $S = \{-4, 0, 3, 7\}$

An extremely versatile way of building sets is through conditional statements. In this situation, we can build finite or infinite sets of all numbers that have certain qualities. Here is an example of a one finite set and one infinite set built with the conditional method:

Example 3.8. $A = \{a \in \mathbb{Z} \mid 0 < a < 10\}$

$$B = \{b \in \mathbb{Z}^+ \mid b \mod 2 = 0\}$$

In this case, set A contains the numbers 1 through 9 and set B contains all even positive integers.

There is one final way of defining a set. This is the constructive method which gives a formula and a set to act upon. An example of this which contains all squared numbers would look like so:

$$C = \{n^2 \mid n \in \mathbb{Z}\}$$

3.3 Relations

In much the same way as we can do relations between conditional statements, we can also relate sets to each others. There are two fundamental set operations, the union and the intersection.

The union of two sets A and B, denoted by $A \cup B$ forms a set such that $A \cup B = \{x \in A \text{ or } x \in B\}$.

The intersection of two sets A and B, denoted by $A \cap B$ forms a set such that $A \cap B = \{x \in A \text{ and } x \in B\}$.

We can string these together to form complex sets like so:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

3.3.1 Equivalence Relations and Partitions

We cannot go any further without explaining the concept of subsets. We call A a subset of B, denoted by $A \subseteq B$ if every element of A is also found in B.
That is to say: if $A \subseteq B$ then $\forall x \in A$, $x \in B$. True equivalence comes when $A \subseteq B$ and $B \subseteq A$ meaning that every element in A is in B and every element of B is in A.

If every element in A exists in B, but there exists some element in B that is not found in A. That is to say:

Definition 3.9. *If $A \subseteq B$ and $\exists x \in B$ such that $x \notin A$ then we call A a proper subset of B and write $A \subset B$*

We can also subtract sets from one another. The set formed by the difference of two sets A and B is defined as follows:

Definition 3.10. $A - B = \{x \mid x \in A$ *and* $x \notin B$

Two sets A and B are called disjoint if there exist no elements in common between A and B. That is to say:

Definition 3.11. $A \cap B = \emptyset$

A partition of X is a subset Z of X which has the following properties:

(i) The subsets in Z are non-empty
(ii) The subsets in Z are disjoint
(iii) The subsets in Z cover all of X.

We can partition sets in any number of ways. For example, we can partition the set of Integers into even and odd numbers.

3.4 Mathematical Induction

Mathematical Induction is one of the most powerful tools in the mathematicians arsenal. It allows one to prove a statement about a base case, prove that this remains true in successive cases, and then imply that it holds for all successive cases.

3.4.1 Principle of Mathematical Induction

The basic principle of mathematical induction can be stated as follows:

Theorem 3.12. *Suppose that P(n) is a statement involving a general positive integer n. Then P(n) is true for all positive integers 1,2,3,... if:*
(i) P(1) is true, and
(ii) P(k) $\Rightarrow P(k+1)$ *for all positive inters k.*

Step 2 is known as the induction step, and we must be very careful in doing our induction step so as ensure that our proof is valid.

Here is a template to help in the induction process provided by Peter Eckles:

Proof: We use induction on n.

Base case: [Prove the statement P(1)]
Inductive step: Suppose now as an inductive hypothesis that [P(k) is true] for

some positive integer k. Then [deduce that P(k+1) is true]. This proves the induction step.
Conclusion: Hence, by induction [P(n) is true] for all positive integers n.

Here is an example of mathematical induction proof:

Example 3.13. *Prove that for all positive integers n the number $n^2 + n$ is even*

Proof: We use induction on n.

Base case: For n = 1, $n^2 + n = 1 + 1 = 2 = 2 \times 1$, an even number, as required.

Inductive step: Suppose now as an inductive hypothesis that $k^2 + k$ is even for some positive integer k. Then $k^2 + k = 2q$ for some $q \in \mathbb{Z}$

Then $(k+1)^2 + (k+1) = k^2 + 2k + 1 + k + 1 = (k^2 + k) + 2k + 2$

$(k^2 + k) + 2k + 2 = 2q + 2k + 2 = 2(q + k + 1) = 2p$, where p is the integer $q + k + 1$, and so $(k+1)^2 + (k+1)$ is even as was required.

Conclusion: Hence, by induction, $n^2 + n$ is even for all positive integers n.

3.5 Functions

Functions at their most basic level, a function is a machine which takes something as input and outputs something else. For our purpose, a function is a process that maps a domain X onto a codomain Y. If a function maps a domain X onto Y, we will write $f : X \to Y$
Given $x \in X$ and $f(x) \in Y$, we may also write $x \mapsto f(x)$ with f(x) being called the image of x under f, X is called the domain, and Y is called the codomain.

There exist two fundamental properties that a function may have which are of interest to us.

The first property concerns whether a function is 1 to 1. This means no two values in the domain map to the same value in the codomain. That is to say that given $x_1 \neq x_2$, we can be assured that $f(x_1) \neq f(x_2)$
This property is also known as an injection.

The second property that a function may have is whether it is onto, in relation to a domain X and Codomain Y. To put it in another way

$$\forall y \in Y, \exists x \in X, y = f(x)$$

A function that has this property is known as a surjection.

A function that is both an injection and a surjection is known as a bijection.

4 Modern Algebra

Now we will be moving on to the field of modern, or abstract algebra. This is a field concerned with the structure, and behavior of systems where the regular rules that we've learned about mathematical operations may not necessarily apply.

Information in this summary is gleaned from [8]. The examples included are also from the same source. My own notes regarding this course were lost.

4.1 Groups and Isomorphisms of Groups

Fundamental to the discussion of modern algebra is the idea of a group. Let us first define a group

Definition 4.1. *A* group $\langle G, * \rangle$ *is defined as a set G closed under a binary operation *, such that several axioms are satisfied. These axioms are:*

Axiom 1: $\forall\ a, b, c \in G$ *we have*
$(a * b) * c = a * (b * c)$.
*That is to say, * is associative in G*

Axiom 2: $\exists e \in G$ *such that* $\forall x \in G$ *we have*
$e * x = x * e = x$.
*In other words, an identity element exists in G for the operation *.*

Axiom 3: Corresponding to each $a \in G$, *there is an element* $a' \in G$ *such that*
$a * a' = a' * a = e$.
*To put it simply, every element in G has an inverse element in G with respect to the operation *.*

Please note that a group is not necessarily *abelian*

Definition 4.2. *A group is called abelian if its binary operation is commutative.*

This is a very big definition, and there are a few bits of housekeeping regarding definitions that we must take care of before we move on. First, since a group is defined as a set closed under a binary operation *, lets talk about binary operations.

Definition 4.3. *A binary operation * on a set S is a function mapping $S \times S$ into S.*
For each (a, b) in $S \times S$, we will denote the element $*((a, b))$ of S by $a * b$

We may regard a binary operation * on a set S as a function assigning an element $a * b$ to each ordered pair (a, b) in S.

We will now proceed with a couple of examples of binary operations to ensure that we are familiar with their properties in action.

Example 4.4. *The normal addition operation and the normal multiplication operation is a binary operation on some of the fundamental sets such as* $\mathbb{R}, \mathbb{Z}, \mathbb{C}$ *et cetera.*

In the previous example, addition was a binary operation because the result of the two operands couldn't be outside the set. New we'll take a look at the addition operation in a different set and see if it is still a binary operation

Example 4.5. *Let + be the usual addition operation and let a set A include all of* $\mathbb{Z}^+$ *and -1.*
If + is a binary operation, then $a + b \in A\ \forall a, b \in A$
However, if $a = b = -1$ *then* $a + b = -2$ *and* $-2 \notin A$.
Therefore, we can conclude that + is not a binary operation on A

This shows us that even an operation which we know to be a binary operation in many cases is not necessarily one for a particular set.

Example 4.6. *Let / be the usual division operation. If / is a binary operation on* $\mathbb{Z}$ *then* $a/b \in \mathbb{Z}$ *for all* $a, b \in \mathbb{Z}$
However, if $a < b$ *then* $a/b \notin \mathbb{Z}$.
Thus we say that / is not a binary operation on $\mathbb{Z}$

This shows us that even a set which we know to have binary operations will not necessarily support a particular binary operation.

Now we will take a look at the particular traits that the operation * may have in a group, namely: Commutativity and Associativity

Theorem 4.7. *Commutativity: A binary operation * on a set S is* commutative *if and only if* $a * b = b * a$ *For all* $(a, b) \in S$

Theorem 4.8. *Associativity: A binary operation * on a set S is* associative *if and only if* $a * (b * c) = (a * b) * c$ *For all* $(a, b, c) \in S$

Now that we have a sense of binary operations, we will take our discussion into the realm of isomorphisms of binary structures. Please note that we are speaking about general binary structures and not about groups in particular which are a subclass.
Let us first work to define an isomorphism:

Definition 4.9. *Let* $\langle S, * \rangle$ *and* $\langle S', *' \rangle$ *be binary algebraic structures.*
An isomorphism *of S with S' is a one to one function* ϕ *mapping S onto S' such that:*
$\phi(x * y) = \phi(x) *' \phi(y)$ *for all* $x, y \in S$.
This is the Homomorphism Property

If such a mapping (ϕ) exists, then we say that S and S' are isomorphic binary structures. There are four steps that go into proving that two structures $\langle S, * \rangle$ and $\langle S', *' \rangle$ are isomorphic.

Step 1: Define our function ϕ which gives us our isomorphism of S to S'.
Step 2: Show that ϕ is a one to one function.
Step 3: Show that ϕ is onto S'
Step 4: Show that $\phi(x * y) - \phi(x) *' \phi(y)$.

Properties of a binary structure can be divided into two separate classes: structural properties and non-structural properties.

Structural properties are those properties that *must* be shared by any isomorphism of the structure. Examples include structure size (if two structures are isomorphic, they must have the same cardinality) or commutativity.

Non-structural properties of a binary structure are those that may or may not be shared by an isomorphism of the structure. Examples of this class include things like having a particular element or the chosen set (i.e. the set $\mathbb{Z}^-$ can be an isomorphism of $\mathbb{Z}^+$ with the addition operation.

Now that we've done all of our housekeeping work, lets move back into groups and show some examples of structures that may or may not be groups.

Example 4.10. *The binary structure $\langle \mathbb{Z}^+, + \rangle$ is not a group because Axioms 2 and 3 of our definition do not hold. There is not inverse element in $\mathbb{Z}^+$ nor is there an identity element.*

Example 4.11. *The usual addition operation with the sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are abelian groups.*

Example 4.12. *The set $\mathbb{Z}^+$ with the usual multiplication operation is not a group because Axiom 3 is violated.*
While there may be an identity element, 1, there is no inverse element for any element but 1.

We now move on to a few interesting theorems regarding the properties of groups.

Theorem 4.13. *If G is a group with binary operation *, then $a * x = b$ has a unique solution x in G for all $a, b \in G$.*

Corollary 4.14. *please note that in addition to this, if the group is non-abelian, then $a * x = b$ and $y * a = b$ have unique solutions x and y for $a, b \in G$*

This theorem is a natural extension of our Axioms and can be proven as follows:

Proof. First we show the existence of at least one solution by just computing that $a' * b$ is a solution of $a * x = b$. Note that:

$$a * (a' * b) = (a * a') * b$$

$$= e * b,$$

$$= b$$

Thus $x = a' * b$ is a solution of $a * x = b *$. We can use a very similar proof to prove our corollary. □

What follows is a theorem regarding the existence and uniqueness of identity elements and inverse elements

Theorem 4.15. *In a group G with binary operation *, there is only one element e in G such that*

$$e * x = x * e = x$$

for all $x \in G$. Likewise, for each $a \in G$, there is only one element a' in G such that

$$a' * a = a * a' = e$$

Proof. For the identity part of this theorem, let us suppose that both e and e' are identity elements of a binary structure S with respect to an operation *. This gives us a system of equation with regards to identity.

$$e * e' = e'$$

$$e' * e = e.$$

This system shows that $e = e'$. That is to say, there is only one identity element.

For the inverse part of the proof, let us suppose that $a \in G$ which has inverses a' and a" such that $a' * a = a * a' = e$ and $a'' * a = a * a'' = e$. Then

$$a * a'' = a * a' = e$$

Thus, by the cancellation laws,

$$a'' = a'$$

. This shows that the inverse of a in a group G is unique. □

Corollary 4.16. *Let G be a group. For all $a, b \in G$, we have $(a * b)' = b' * a'$.*

Theorem 4.17. *If G is a group with binary operation *, then the left and right cancellation laws hold in G, that is, $a*b = a*c$ implies that $b = c$, and $b*a = c*a$ implies that $b = c$ for all $a, b, c \in G$*

Proof. Suppose $a * b = a * c$ then by the third group axiom there exists a', and

$$a' * (a * b) = a' * (a * c)$$

thus $b = c$
We can use a similar proof with $b * a = c * a$ to prove the other side of the cancellation law □

It is interesting to not that there also exists some binary structures with weaker strictures than a group.

They come in a couple of different flavours including the semigroup, a set with an associative binary operation; another type is a monoid which is a semigroup with an identity element.

All that it lacks is an inverse element which differentiates it from a group.

Last but not least in our discussion of groups, we'll be talking about finite groups. Up until now, we've focused on our attention on infinite groups, that is to say, groups with an unlimited set.

Obviously, the empty set cannot give rise to a group because it has no elements on which to do a binary operation.
This gives rise to an interesting question: what is the minimal set which may give rise to a group?
as it turns out, the minimal set which can give rise to a group is the set containing only the identity element. That is to say $\{e\}$.

Example 4.18. *Let the set S be the set 1, this is a group under the usual multiplication operation.*

Let us look at a larger group S which consists of the elements e and a. We can put all the possible binary operations into the following table:

*	e	a
e		
a		

We see here that we must define the result of $a * a$, if we define it as a or e, the result will be a group.

*	e	a
e	e	a
a	a	e

If the result is something outside the bounds of our finite set, we will not have a group.

5 History of Mathematics

Mathematics has a long and storied history. Every society on earth has made their own mathematics and contributed to the state of the art today. Volumes and volumes have been written that take us from the beginning mathematics to the modern day. Today however, we will focus very closely on the sizeable contributions of ancient greece to the field of mathematics

Information in this summary is gleaned from [1]. Please note that this summary is drawn in large part from a paper originally written for the history of math course offered at Indiana University East.

5.1 Overview of Ancient Greek Mathematics

The contributions of Greek mathematicians to the body of knowledge concerning mathematics are impossible to overstate. They founded and formalized the idea of proof, which is so central to mathematics today. They also made massive strides in the field of geometry, which will be a large part of our focus today. Additionally they made major advances in number theory and even came close to stumbling upon some of the ideas of calculus. Because of this, it is hard to find any field of mathematics which isnt in some way influenced by the Greeks

Mathematics in Ancient Greece can be largely divided into two broad categories based upon time.

This period of time spans from around 600 BCE to around 146 BCE. There are literally dozens of eminent mathematicians and philosophers hailing from Greece during this period of time, but in the interest of brevity, well be limiting our discussion today to six giants among giants: Thales, Pythagoras, Aristotle, Euclid, Archimedes and Eratosthenes. Well be taking our study of these men in chronological order so as to better understand the developments and contributions of Greek math.

5.1.1 Thales of Miletus

Thales of Miletus is often touted as the first Greek mathematician and moreover, the first true scientist of the west. As such, it would seem fitting to start our story here. Thales was born In 624 BCE and lived into his eighties, passing away in 546 BCE. Much of our knowledge of Thales is anecdotal and through the writings of others. What is certain is that he had an advanced (for the time) conception of geometry and applied it to great effect. One of the early accomplishments attributed to Thales was the prediction of a solar eclipse in 585 BCE. Given that there was no observable cycle for solar eclipses, this is quite an accomplishment, though there is no record of how he came to predict it. Another tale of Thales brilliance comes in the account of how he was able to measure the heights of the pyramids to a reasonable degree of accuracy. He did it by measuring the length of the shadow the pyramid cast when the shadow he cast was precisely as tall as he was. This achievement implies a body of knowledge that was peerless in its time. First and most importantly it means that he understood the idea of similar triangles, meaning that he recognized that when him and his shadow formed an equilateral triangle, so too must the pyramid and its shadow. This leads us neatly into the nitty-gritty of his actual contributions to the field. There is some contention on the issue, but it is debated that Thales should be credited with the following 5 theorems of elementary geometry.

Theorem 5.1. *A circle is bisected by its diameter*

This theorem is basic and quite handy as it gives us definite dimensions on any line which bisects a circle.

Theorem 5.2. *The base angles of an isosceles triangle are equal*

This demonstrates a more advanced understanding of the qualities of triangles than was common at the time.

Theorem 5.3. *The angles between two intersecting straight lines are equal*

Theorem 5.4. *Two triangles are congruent if you have two and angles and one side equal to each other*

This again demonstrates an understanding of trigonometry that was advanced for the time and extremely useful.

Theorem 5.5. *An angle in a semicircle is a right angles*

This list of theorems is absolutely fundamental to geometry and forms part of many other works and proof that proceed it. The last of these is fittingly called Thales theorem. Of course, he also had a few more fantastic ideas which were less rooted in fact and rigor than speculation. He believed that earthquakes were the result of the earth being rocked by waves as it floated in a vast ocean, and further believed that water was the irreducible substance from which all things were formed. While this was incorrect, it was also the first recorded attempt to describe the materiel world in terms of a small, finite number of different elements.

5.1.2 Pythagoras

Pythagoras was an incredible, but somewhat odd mathematician who lived from 569 BCE to 475 BCE, having some overlap with Thales. Of course, when we think of Pythagoras, our minds instantly jump to the theorem known by children the world over

$$a^2 + b^2 = c^2$$

with a and b being the length of the sides of a right triangle, and c being the length of the hypotenuse. Unfortunately for Pythagoras, the theorem attributed to him was discovered by the Babylonians a millennium earlier; however, it is possible that he may have been the first to rigorously prove it.

His most important innovation was something that today we take almost completely for granted: the idea of abstraction. Before Pythagoras, mathematics was set in the concrete; with numbers being used only as a representation of actual quantities. In a sense, this is part of why so much of early math is concerned with geometry. In a sense, Pythagoras invented the field of pure math by way of divorcing numbers from the limited domain of representation, and allowed them to be objects in their own right. This concept of abstraction allowed the idea of abstract mathematical proof to flourish. His fascination with numbers as objects arguably founded number theory, often touted as the queen of mathematics.

His love of music led him to attempt to codify music with mathematical formulae, which allowed him to discover the ratios of the keys and scales. He also believed that the motion of the planets was inherently quantifiable and formulaic, and related it to his work with music in what he called The music of the spheres.

He also created an order known, appropriately enough, as the Pythagoreans. This group was one part academic, another part, monastic order, and another part of something akin to a cult. The tenets of the order are a bit hazy, but we do know that they held an almost religious reverence to the mystical math that they worked at. They also may have believed in the inherent magical qualities of certain numbers and had a miscellany of interesting beliefs regarding ritual. One very fascinating feature of the group was a belief that reality was inherently quantifiable, which is certainly a progressive thought for the time. But now we must jump forward by more than a century and take a look at Aristotle.

5.1.3 Aristotle

Aristotle presents us with a problem, since he was known primarily as a philosopher rather than a mathematician. As a matter of fact, most if not all of the great Greek mathematicians were considered philosophers, although the term carried a broader meaning than it does today.

Aristotles contributions to math are more general than the other people on our list. His contribution to math lies in logic, without which mathematics could not exist. He codified term logic, which was extremely influential and useful in every avenue of science.

He is also credited with codifying the 3-line syllogistic argument, which is an excellent foundational idea for proof through relation. Now we move on to one of the most prolific luminaries on our list today: Euclid.

5.1.4 Euclid

Euclids life is an enigma; almost all we know of him can be summed up in this sentence: He was born sometime in the mid-4th century BCE, taught in Alexandria, and died somewhere towards the middle of the 3rd century BCE.

Euclid made sizeable and fantastic contributions to the field of geometry. For Euclid, the study begins and ends with his book *Elements*. Euclids *Elements* was a series of 13 books written on the subject of math and geometry, which is widely considered to be the most influential textbook in the history of the written word. The book works from a series of basic axioms, definitions, and postulates and derives from there.

It touches upon geometric algebra, proportion, spatial geometry, and even number theory; acting much like an ark, ferrying the sum of mathematical knowledge at the time to us in the contemporary world. The importance of this work is impossible to overstate, and while we could continue to trace its influence upon the mathematics, we must now move on to the eminently practical Archimedes.

5.1.5 Archimedes

Archimedes is one of the most spectacular mathematicians of history. Archimedes spent much of his life wearing the hat of engineer rather than mathematician, but mathematics was where his heart truly resided. His achievements span geometry and he even touched upon the ideas of calculus, which would be codified over 1600 years after his death in 212 BCE. He used the method of exhaustion and even an early form of limits to prove many geometric theorems and relationships and found results including an approximation of pi which are almost unreal in their accuracy. Many of his achievement are more physical in nature.

He invented the Archimedes screw, which is a device utilizing a spiral (which was a near obsession for Archimedes) to efficiently lift water. He also codified the Archimedes principle which states that the buoyant force exerted on an object in fluid is equal to the weight of the fluid displaced. There is an apocryphal story about Archimedes which relates to water and displacement.

In short, a crown was made for a temple with the instructions to use pure gold in its construction, but it was suspected that the smiths might have alloyed the gold with a lesser metal to save on production costs. Archimedes was asked to figure out whether such subterfuge had taken place without damaging the crown.

Had it been a cube or some regular shape, it would have been easy as he could simply use volume and weight to determine density. Archimedes was stumped and went to take a bath. He observed that the water he displaced was equal to his volume.

He realized that he could use this to determine the density of the crown and in his excitement he took to the streets stark naked, shouting Eureka! meaning I have found it. Unfortunately, Archimedes of Syracuse was killed when the Romans sacked his city.

The story goes that he was drawing circles in the dirt, working out some problem. A Roman soldier grew impatient and struck him down with his last word being something to the effect of dont disturb my circles.

5.1.6 Eratosthenes

Last but not least we come to Eratosthenes of Cyrene. Eratosthenes main accomplishments come in the disparate fields of geography and number theory. His best known accomplishments involve his derivation of the circumference of the earth and the earths axial tilt both of which he determined to a high degree of accuracy. His less widely known but also extremely important accomplishment was the invention of the so called Sieve of Eratosthenes which is an extremely efficient algorithm for finding prime numbers still in use today. The sieve works in an incredibly simple manner. The algorithm works quite simply as follows:

Theorem 5.6. *Sieve of Eratosthenes*

First make a list of every number from 2 to some arbitrary end point n.
Then cross out every second number (4,6,8, etc) after 2.
Then cross out every third number from three, and every fifth from five.
Repeat until you hit $\frac{n}{2}$.
The numbers remaining are all the prime numbers in that span.

This allows you to find arbitrarily large prime numbers with little computation.

However, it grows less efficient when the end point gets extremely large since it means you have to run through your number line many times. This simple algorithm is a great accomplishment which at the time may have been esoteric.

After all, a theorem regarding prime numbers does not seem to have an immediate application. However, today prime numbers form the basis for cryptography and computer security. It goes to show that even the most exotic of results may eventually find an application.

6 Mathematical Modeling

Mathematics is a precise, idealistic and pristine world. By contrast, the physical universe is a thing of complexity and relative randomness. In mathematical modelling we work to approximate real world phenomena with mathematical models. The advantages of this approach for projection, prediction, and analysis are enormous.

Please note that unless otherwise stated, the information in following summery of mathematical modelling are gleaned from [4]. Further note that the examples are drawn from the text as well, as my own personal notes have been lost.

6.1 Simplex Method

The simplex method is a linear programming method developed by George Dantzig which incorporates test of both optimality and feasibility in seeking out the optimal solution to a linear program, provided a solution exists.

Steps in the Simplex Method

1. Tableau Format: Place the linear program into Tableau Format, which will be explained after this list.

2. Initial Extreme Point: The Simplex Method begins with a known extreme point, usually the origin $(0, 0)$

3. Optimality Test: Determine if an adjacent intersection point improves the value of the object function. If not, the current extreme point is optimal. If an improvement is possible, the optimality test determines which variable currently in the independent set (having value 0) should enter the dependent set and become non-zero.

4. Feasibility Test: To find a new intersection point, one of the variables in the dependent set must exit to allow the entering variable from step 3 to become dependent (only one variable may be independent at a time). The feasibility test determines which variable to choose for exiting ensuring feasibility.

5. Pivot: Form a new equivalent system of equations by eliminating the new dependent variable from the equations that do not contain the variable from step 4. Then set the new independent variables to 0 and the new system to find the values of the new dependent variables, thereby determining an intersection point.

6. Repeat: Repeat steps 3-5 until an optimal extreme point is found.

This wall of text, while helpful as a guide does little to enlighten us on the method's use. In order to shed some more light on the simplex method, lets examine an example from the text.

Example 6.1. *Maximize the function*

$$25x_1 + 30x_2$$

subject to

$$20x_1 + 30x_2 \leq 690$$

$$5x_1 + 4x_2 \leq 120$$

$$x_1, x_2 \geq 0$$

Step 1: Tableau Form
First we must adjoin a new constraint to ensure that any solution improves on our best found solution thus far. As a result, we will add in the restraint

$$25x_1 + 30x_2 \geq 0$$

Since all of our constraints must be $\leq$ inequalities for the method to work, we will multiply by -1 in order to get the following complete constraint set:

$$20x_1 + 30x_2 \leq 690$$

$$5x_1 + 4x_2 \leq 120$$

$$-25x_1 - 30x_2 \leq 0$$

It is worth noting that the Simplex Method assumes that all variables will only assume non-negative values, so the constraint $x_1, x_2 \geq 0$ does not need to be repeated from this point forwards.

Next we need to turn the inequalities of our constraints into equalities by the addition of a new non-negative variable to each of our constraints. This variable is called a slack variable *because it is a measure of how much slack there is in our values; by this we mean, how close our get us to the bounds of our constraints. In simple, ideal cases studied in classrooms, we will often see the slack variables taking on a values of 0 in the end, as we end up with solutions that are at the very edges of our constraints. In any case, after adding in our slack variables, we end up with the following* augmented constraint set

$$20x_1 + 30x_2 + y_1 = 690$$

$$5x_1 + 4x_2 + y_2 = 120$$

$$-25x_1 - 30x_2 + z = 0$$

If you look closely, you'll realize that z is in fact the value of the objective function we are looking to optimize.

Step 2: Initial Extreme Point
Because there are two decision variables, all possible intersection points lie in the x_1, x_2 plane and can be determined by setting two of the variables $\{x_1, x_2, y_1, y_2\}$

to zero. (The variable z is always a dependent variable because it represents the value of the objective function at the extreme point in question.) The origin is a feasible extreme point and corresponds to the extreme point characterized by $x_1 = x_2 = 0$, $y_1 = 690$, *and* $y_2 = 120$. *Thus,* x_1 *and* x_2 *are independent values whose values are set to 0 and* y_1, y_2, *and z are dependent variables whose values are then determined. As we previously mentioned, z conveniently records the current value of the objective function at the extreme points of the convex set in the* x_1, x_2 *plane as we compute them by elimination.*

Step 3: Optimality Test

In the preceding format, a negative coefficient in the last (or objective function) equation indicates that the corresponding variable could improve the current objective function value. Thus, the coefficients -25 *and* -30 *indicate that either* x_1 *or* x_2 *could enter and improve the current objective function value of* $z = 0$ *(The current constraint corresponds to* $z = 25x_1 + 30x_2 \geq 0$ *with* x_1 *and* x_2 *currently independent and 0.) When more than one candidate exists for the entering variable, a rule of thumb for selecting the variable to enter the dependent set is to select that variable which has the largest (in absolute value) negative coefficient in the objective function row. If no negative coefficients exists, then the current solution is optimal. In the case at hand, we choose* x_2 *as the new entering variable. (The procedure is inexact at this stage because we do not know what values the entering variable can assume.)*

Step 4: Feasibility Test

The entering variable x_2 *(in our example) must replace either* y_1 *or* y_2 *as a dependent variable (because z always remains the third dependent variable). To determine which of these variables is to exit the dependent set, first divide the right hand side values 690 and 120 (associated with the original constraint inequalities) by the components for the entering variable in each inequality (30 and 4, respectively, in our example) to obtain the ratios* $\frac{690}{30} = 23$ *and* $\frac{120}{4} = 30$. *From the subset of the ratios that are positive (both in this case), the variable corresponding to the minimum ratio is chosen for replacement (*y_1 *corresponding to 23 in this case). The ratios represent the value the entering variable would obtain if the corresponding exiting variable were assigned the value 0. Thus, only positive values are considered and the smallest positive value is chosen so as not to drive an variable negative. For instance, if* y_2 *were chosen as the exiting variable and assigned the value 0, then* x_2 *would assume the value 30 as the new dependent variable. However, then* y_1 *would be negative, indicating that the intersection point* $(0, 30)$ *does not satisfy the first constraint. The* ***minimum positive ratio rule*** *illustrated previously obviates the enumeration of any infeasible intersection points. In the case at hand, the dependent variable corresponding to the smallest ratio, 23 is* y_1, *as such it becomes the exiting variable. Thus* x_2, y_2, *and z form the new set of dependent variables and* x_1 *and* y_1 *form the new set of independent variables.*

Step 5: Pivot

Next we derive a new (equivalent) system of equations by eliminating the entering variable x_2 in all the equations of the previous system that do not contain the exiting variable y_1. After the elimination, we then find the values of the dependent variables x_2, y_2, and z when the independent variables x_1 and y_1 are assigned the value 0 in the new system of of equations. This is called the ***pivoting procedure.*** *The values of x_1 and x_2 give the new extreme point (x_1, x_2) and z is the new and improved value of the objective function at that point.*

Step 6: Repeat until satisfied
After performing the pivot, the optimality test is applied again to determine if another candidate entering variable exists. If so, we choose an appropriate one and apply the feasibility test again to choose an exiting variable. Then the pivoting procedure is performed again. The process is repeated until no variable has a negative coefficient in the objective function row. At this point we have found our optimal extreme point.

Now that we have the idea of the Simplex method thoroughly fleshed out in relation to our problem. Let us revisit the problem using tableaus in order to better see the operations taking place.

Example 6.2. *Tableau Problem:*

Tableau 0 (Original Tableau)

x_1	x_2	y_1	y_2	z	*RHS*
20	*30*	*1*	*0*	*0*	$690(= y_1)$
5	*4*	*0*	*1*	*0*	$120(= y_2)$
-25	*-30*	*0*	*0*	*1*	$0(= z)$

Dependent variables: $\{y_1, y_2, z\}$
Independent variables: $x_1 = x_2 = 0$
Extreme Point: $(x_1, x_2) = (0, 0)$
Value of Objective Function: $z = 0$

Optimality Test: *The entering variable is x_2 (corresponding to -30 in the last row).*

Feasibility Test: *Compute the ratios for the RHS to determine minimum positive ratio*

x_1	x_2	y_1	y_2	z	*RHS*	*Ratio*
20	*30*	*1*	*0*	*0*	$690(= y_1)$	$23(= 690/30)$
5	*4*	*0*	*1*	*0*	$120(= y_2)$	$30(= 120/4)$
-25	*-30*	*0*	*0*	*1*	$0(= z)$	$*$

So we choose y_1 corresponding to the minimum positive ratio 23 as the exiting variable.

Pivot: *Divide the row containing the exiting variable by the coefficient of the entering variable in that row giving a coefficient of 1 for the entering variable in this row. Then eliminate the entering variable x_2 from the remaining rows (which do not contain the exiting variable y_1 and have a 0 coefficient for it.) The results are summarized in the next tableau.*

Tableau 1

x_1	x_2	y_1	y_2	z	*RHS*
.66667	*1*	*.03333*	*0*	*0*	$23 (= x_2)$
2.3333	*0*	*-.1333*	*1*	*0*	$28 (= y_2)$
-5	*0*	*1*	*0*	*1*	$690 (= z)$

Dependent variables: $\{x_2, y_2, z\}$
Independent variables: $x_1 = y_1 = 0$
Extreme Point: $(x_1, x_2) = (0, 23)$
Value of Objective Function: $z = 690$

The pivot determines that the new dependent variables have the values $x_2 = 23,\ y_2 = 28,$ *and* $z = 690$

Optimality Test: *The entering variable is x_1 (corresponding to -5 in the last row).*

Feasibility Test: *Compute the ratios for the RHS to determine minimum positive ratio*

x_1	x_2	y_1	y_2	z	*RHS*	*Ratio*
.66667	*1*	*.03333*	*0*	*0*	23	$23 (= 34.5/.66667)$
2.33333	*0*	*-.13333*	*1*	*0*	28	$12 (= 28/2.33333)$
-5	*0*	*1*	*0*	*1*	690	$*$

Choose y_2 as the exiting variable because it corresponds to the minimum positive ratio 12.

Pivot: *Divide the row containing the exiting variable (the second row in this case) by the coefficient of the entering variable in that row (the coefficient x_1 in this case) giving a coefficient of 1 for the entering variable in this row. Then eliminate the entering variable from the remaining rows (which do not contain the exiting variable and have a 0 coefficient for it). The results are summarized* *in our last tableau.*

Tableau 2

x_1	x_2	y_1	y_2	z	*RHS*
0	*1*	*.071429*	*-.28571*	*0*	$15(= x_2)$
1	*0*	*-.057143*	*.42857*	*0*	$12(= x_1)$
0	*0*	*.714286*	*2.14286*	*1*	$750(= z)$

Dependent variables: $\{x_2, x_1, z\}$
Independent variables: $y_1 = y_2 = 0$
Extreme Point: $(x_1, x_2) = (12, 15)$
Value of Objective Function: $z = 750$

Optimality Test: *There are no negative coefficients in the bottom row, thus* $x_1 = 12$ *and* $x_2 = 15$ *gives the optimal solution of* $z = 750$ *for the objective function.*

Please note that while there were 6 intersection points in our constraints, we needed only to enumerate two of them to find our optimal solution. This is a big part of the draw of the simplex method. It can reduce the number of computations required to find the optimal extreme point of a linear program. This comes in extremely handy when doing very large linear programs with many possible intersections.

7 Differential Equations

Differential Equations are a tremendously powerful tool in the mathematicians arsenal. In particular, the utility of differential equations in modelling systems is almost boundless. The information for the following summary of Differential Equations is taken from [2]. Please note that the examples presented in the following summary section are also provided in the text as I no longer have notes from a Differential Equations course.

7.1 Ordinary Differential Equations

To put it simply, a differential equation is one in which derivatives are used as variables in an equation. In the simplest case, we have what is called an *ordinary differential equation.*

Definition 7.1. *An* ordinary differential equation *is an equation relating an unknown function of one variable to one or more functions of its derivatives.*

Here are some examples of ordinary differential equations with x as our unknown as a function of t. That is to say $x = x(t)$.

Example 7.2.

$$\frac{dx}{dt} = t^7 \cos x$$

$$\frac{d^2x}{dt^2} = x\frac{dx}{dt}$$

$$\frac{d^4x}{dt^4} = -5x^5$$

Each differential equation in this example is of a different *order.*

Definition 7.3. *The* order *of a differential equation is the order of the highest derivative of the unknown function (dependent variable) that appears in the equation.*

In our example set of differential equations, the orders are first, second, and fourth order respectively.

Very often, our dependent variable is a function of time, denoted as t. Applications for this type of equation include

1. Population Dynamics
2. Mixture and Flow Problems
3. Electronic Circuits
4. Mechanical Vibrations and Systems

We will now show an example problem regarding population dynamics, which should give us a good idea of the type of problems differential equations hopes to answer.

Example 7.4. *Population Dynamics Over Time*

Assume that the population of Washington, DC, grows due to births and deaths at the rate of 2% *per year and there is a net migration into the city of* 15000 *people per year. Write a mathematical equation that describes the situation.*

SOLUTION. We let

$$x(t) = \textit{population as a function of time } t.$$

From Calculus,

$$\tfrac{dx}{dt} = \textit{rate of change of the population } x$$

In our example, 2% *growth means* 2% *of the population* $x(t)$*. Thus, the population of Washington, DC satisfies the following*

$$\frac{dx}{dt} = 0.02x + 15000.$$

Physical laws often lead to differential equations which are quite handy. In our next example, we will look at the applications of differential equations to free-falling bodies in open air.

Example 7.5. *The Joys of Free-fall*

Newton's Law says

$$F = ma$$

where m *is constant mass, and* a *is acceleration, the second derivative of position. This allows us to rewrite Newton's Law as*

$$m\frac{d^2x}{dt^2} = F.$$

If the forces acting on a particular body going into free-fall are gravity (which we define as $-mg$ *and the force of air resistance, which we know to be proportional to velocity (*$\frac{dx}{dt}$*, then our position at any given time satisfies the following second order differential equation*

$$m\frac{d^2x}{dt^2} = -mg - c\frac{dx}{dt}$$

. This is a bit underwhelming as a result. However, suppose we are tied to a spring which exerts an additional force proportional to our position that satisfies

Hooke's Law. In this case, we will end up with the following, more interesting second order differential equation

$$m\frac{d^2x}{dt^2} = -mg - kx - c\frac{dx}{dt}$$

Many of the physical laws concerning electricity give rise to differential equations as well. As such, we can model the behaviour of circuits with differential equations quite handily. For example, a simple RLC circuit can be modelled by the following second-order differential equation:

$$L\frac{d^2i}{dt^2} + R\frac{di}{dt} + \frac{1}{C}i = f(t)$$

But if we derive or are given a possible solution to a differential equation, how can we verify that what we have is actually a solution?

Example 7.6. *Showing that a Function is a Solution*
Verify that $x = 3e^{t^2}$ is a solution of the following first-order differential equation

$$\frac{dx}{dt} = 2tx$$

We substitute $x = 3e^{t^2}$ in both the left and right hand sides of our equation. this gives us

$$\frac{d}{dt}(3e^{t^2}) = 6te^{t^2}.$$

Simplifying the left hand side, we end up with

$$6te^{t^2} = 6te^{t^2}.$$

This holds for all t, thus we can see that $x = 3e^{t^2}$ is a solution for our differential equation.

7.1.1 Initial Value Problems

The simplest first-order differential equation possible arises if the function $f(x,t)$ in $\frac{dx}{dt} = f(x,t)$ does not depend on the unknown solution giving us the differential equation

$$\frac{dx}{dt} = f(t).$$

Luckily, the solution to such a system is quite apparent. We can find a solution by way of integration.

However, we can see that integrating like this will give us an arbitrary constant. From this, we can see that differential equations can have an infinite number of possible solutions. A solution like this is called the *general solution* of a differential equation because it is a sort of formula which gives all possible solutions.

In the following example, we will solve a differential equation this way, and toss in the wrinkle of initial values to find a definite solution.

Example 7.7. *Introducing Initial Values*

Consider the following simple differential equation,

$$\frac{dx}{dt} = t^2.$$

By integration, we obtain

$$x = \frac{1}{3}t^3 + c$$

where c is an arbitrary constant. Without further information, this would be where the problem would end.

However, suppose that we are given the information $x = 7$ at $t = 2$, that is to say, $x(2) = 7$. Applying this information to our general solution we get

$$7 = \frac{8}{3} + c.$$

We can then determine that $c = \frac{13}{3}$. Now we have a unique solution to our differential equation which satisfies our initial condition:

$$x = \frac{1}{3}t^3 + \frac{13}{3}.$$

If our equation is packaged to allow this kind of simple integration, then everything is good. However, suppose that we have a situation where using an indefinite integral is not appropriate. In this case, we can use a definite integral. If both sides of the differential equation $\frac{dx}{dt} = f(t)$ are integrated with respect to t from t_0 to t, this gives us

$$\int_0^1 \frac{dx}{d\bar{t}} d\bar{t} = \int_0^1 f(\bar{t}) d\bar{t}.$$

Note that we introduced a dummy variable and are no longer in t. By working out both sides and cancelling, this gives us

$$x(t) = x(t_0) + \int_{t_0}^{t} f(t)dt.$$

Now lets look at an example of this operation in action.

Example 7.8. *Using Indefinite Integrals*

Solve the differential equation

$$\frac{dx}{dt} = e^{-t^2}.$$

Subject to the initial condition $x(3) = 7$.

The function e^{-t^2} does note have an explicit antiderivative. Thus, we will solve using the definite integration from $t_0 = 3$ to $t = 7$. Using our previously found solution, we find that the solution to this initial value problem is

$$x = 7 + \int_3^t e^{-t^2} d\bar{t}.$$

We will end our discussion of differential equations with a very practical example from the area of mechanics.

Example 7.9. *Brake or Break*

There is a car going 76 meters per second down a road when the driver sees a deer down the road. The brakes are applied hard at $t = 2$. Due to inefficiencies in the braking system, acceleration is known to conform to $a = -12t^2$. How far does the car travel after the breaks are applied?

Our differential equation is

$$\frac{d^2x}{dt^2} = -12t^2$$

By integrating like we did in an earlier example we obtain

$$\frac{dx}{dt} = -4t^3 + c_1$$

Integrating again

$$x = -t^4 + c_1 t + c_2$$

Now we go back and apply our initial conditions to find a definite solution. We know that $x = 0$ and $\frac{dx}{dt} = 76$ at time $t = 2$.

By applying our velocity constraint we get

$$76 = -4t^3 + c_1$$

showing us that $c_1 = 108$. After this, we apply our position constraint in order to see

$$0 = -2^4 + 108t + c_2.$$

From this we get that $c_2 = -200$.

We now solve the following equation

$$\frac{dx}{dt} = -4t^3 + 108 = 0$$

to find that our stopping time is $t = 3$ Substituting $t = 3$ this back into our equation $x = -t^4 + 108t - 200$ giving us a stopping distance of 43 meters.

8 Real Analysis

In the collegiate math student's career, Real Analysis is often the course where a student "hits the wall". It is a daunting subject which gives us many indispensable tools for further study and proof.

The following summary of Real Analysis draws its information from[11] unless otherwise stated. Its examples are also drawn from the same source.

8.1 The Real Number System

The Real Number System is extremely inclusive and contains meaningful subsets within it. We follow with a brief overview of some of the most familiar.

- $\mathbb{N}$ the set of natural numbers, the numbers used for counting, defined as

$$\mathbb{N} := \{1, 2, 3...\}$$

- $\mathbb{Z}$ the set of integers, or whole numbers, defined as

$$\mathbb{Z} := \{... - 2, -1, 0, 1, 2, ...\}$$

  For the curious folks, the Z symbol comes from the German *Zahlen* which means numbers.

- $\mathbb{Q}$ the set of rational numbers, or fractions, defined as

$$\mathbb{Q} := \{\frac{m}{n} \in \mathbb{Z}, n \neq 0\}$$

- $\mathbb{Q}^c$ the set of irrational numbers, or numbers which can't be expressed as fractions, defined as

$$\mathbb{Q}^c := \mathbb{R} - \mathbb{Q}.$$

It is worth noting that each of these sets (save for the irrationals) are proper subsets of the preceding set. By this we mean that

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}.$$

8.1.1 Ordered Field Axioms

We now see how expansive the field of real numbers are and what a big task we have before ourselves. Luckily, we can define our system with two sets of Axioms which make our lives much simpler.

Definition 8.1. ***Field Axioms***
There are functions $+$ and $*$ defined on $\mathbb{R}^2 := \mathbb{R}\times\mathbb{R}$, which satisfies the following properties for every $a, b, c \in \mathbb{R}$

- ***Closure Properties:*** *$a + b$ and $a * b$ belong to $\mathbb{R}$.*

- ***Associative Properties:*** *$a+(b+c) = (a+b)+c$ and $a*(b*c) = (a*b)*c$.*

- ***Commutative Properties:*** *$a + b = b + a$ and $a * b = b * a$.*

- ***Distributive Law:*** *$a * (b + c) = a * b + a * c$.*

- ***Existence of the Additive Identity:*** *There is a unique element $0 \in \mathbb{R}$ such that $0 + a = a$ for all $a \in \mathbb{R}$.*

- ***Existence of the Multiplicative Identity:*** *There is a unique element $1 \in \mathbb{R}$ such that $1 \neq 0$ and $1 * a = a$ for all $a \in \mathbb{R}$.*

- ***Existence of Additive Inverses:*** *For every $x \in \mathbb{R}$ there is a unique element $-x \in \mathbb{R}$ such that*

$$x + (-x) = 0.$$

- ***Existence of Multiplicative Inverses:*** *For every $x \in \mathbb{R}$ where $x \neq 0$ there is a unique element $x^{-1} \in \mathbb{R}$ such that*

$$x * (x^{-1}) = 1.$$

Note that with these Field Axioms, we can well define the algebraic laws of the $\mathbb{R}$ system. However, we are well served by having a set of order axioms as well which defines the ordered nature of the real number system. These will seem largely self explanatory. However, they become extremely critical when in the depths of a proof.

Indeed, it often seems like the *intuitive* or *obvious* ideas are the ones that we are most likely to stumble over.

Definition 8.2. ***Order Axioms***
There is a relation $<$ on $\mathbb{R} \times \mathbb{R}$ that has the following properties:

- ***Trichotomy Property:*** *Given $a, b \in \mathbb{R}$, one and only one of the following statements holds:*

$$a < b,\ b < a,\ a = b$$

- ***Transitive Property:*** *For all $a, b, c \in \mathbb{R}$,*

$$a < b \text{ and } b < c \text{ implies } a < c$$

- ***The Additive Property:*** *For all $a, b, c \in \mathbb{R}$,*

$$a < b \text{ and } c \in \mathbb{R} \text{ implies } a + c < b + c.$$

- ***The Multiplicative Properties:*** *For all $a, b, c \in \mathbb{R}$,*

$$a < b \text{ and } c > 0 \text{ implies } ac < bc$$

  and

$$a < b \text{ and } c < 0 \text{ implies } bc < ac.$$

The more eagle-eyed readers will notice that we used $>$, which is a symbol which we have not defined. By $a > b$ we mean that $b < a$.

8.1.2 The Completeness Axiom

We now work towards the third axiom needed in order to define $\mathbb{R}$. However, first we will need two definitions which we refer back to frequently.

Definition 8.3. ***Bounded Above:*** *The set $E \subset \mathbb{R}$ is said to be* bounded above *if and only if there exists some $M \in \mathbb{R}$ such that $a \leq M$ for all $a \in E$, in which case we can call M the* upper bound *of E.*

Note that bounding above does not mean that the upper bound is actually in the set. For example, we could say that the set of integers between 1 and 10 is bounded by 10 million.

Definition 8.4. ***Supremum:*** *A number s is called a* supremum *of the set E if and only if s is an upper bound of E and $s \leq M$ for all upper bounds M of E. (in this case we can say that E has a* finite supremum *and write $s := \sup E$*

Note that while upper bounds are not unique, there is a unique and singular supremum for any set which has a supremum. In the plainest English, the supremum of a set is simply the number larger than all other elements in the set.

There exists a logically equivalent set of terminology for dealing with the lower end of a set which we will define as follows.

Definition 8.5. ***Bounded Below:*** *The set $E \subset \mathbb{R}$ is said to be* bounded below *if and only if there exists some $m \in \mathbb{R}$ such that $a \geq m$ for all $a \in E$, in which case we can call m the* lower bound *of E.*

As with upper bounds, a set may have infinitely many lower bounds.

Definition 8.6. ***Infimum:*** *A number t is called a* infimum *of the set E if and only if t is a lower bound of E and $t \geq m$ for all lower bounds m of E. (in this case we can say that E has a* finite infimum *and write $t := \inf E$*

Our next theorem, which derives from the properties of suprema shows that we may approximate the supremum of a bounded set utilizing points in the bounded set.

Theorem 8.7. ***Approximation Property of Suprema***
If E has a finite supremum and $\varepsilon > 0$ is any positive number, then there exists a point $a \in E$ such that

$$\sup E - \varepsilon < a \leq \sup E$$

Visually, we can imagine our 'a' being squeezed between the lower inequality and upper inequality allowing us to approximate the supremum.

Finally we have the knowledge necessary to appreciate our third fundamental axiom

Definition 8.8. ***Completeness Axiom:***
If E is a non-empty subset of $\mathbb{R}$ that is bounded above, then E has a finite supremum.

The completeness axiom is utilized in many times throughout Real Analysis. As an example, we can use it to prove the following theorem.

Theorem 8.9. ***Archimedean Principle:***
Given real numbers a and b, with $a > 0$, there is an integer $n \in N$ such that $b < na$

Proof. If $b < a$, set $n = 1$. If $a \leq b$, consider the set $E = \{k \in \mathbb{N} : ka \leq b\}$. E is non-empty since $1 \in E$. Let $k \in E$. Since $a > 0$, it follows that $k \leq \frac{b}{a}$. This proves that E is bounded above by $\frac{b}{a}$. Thus, by the completeness axiom, E has a finite supremum s that belongs to E, in particular, $s \in \mathbb{N}$.
Set $n = s + 1$. Then $n \in \mathbb{N}$ and n cannot belong to E. Thus $na > b$. □

We use the archimedean principle in the proof of our next theorem. This can give you an idea of the incremental, building nature of Real Analysis. By this we mean, we prove one thing, and use that to prove the next thing, and so on.

Theorem 8.10. ***Density of Rationals***
If $a, b \in \mathbb{R}$ satisfy $a < b$, then there is a $q \in \mathbb{Q}$ such that $a < q < b$

In laymans terms, no matter how closely two rationals lie to each other, there will always be a rational number between them.

Proof. Suppose first that $a > 0$. Since $b - a > 0$, we may use the Archimedean Principle to choose a natural number n such that

$$n > \max\{\frac{1}{a}, \frac{1}{b-a}$$

and observe that both $\frac{1}{n} < a$ and $\frac{1}{n} < b - a$.

Consider the set $E = \{k \in \mathbb{N} : \frac{k}{n} \leq a\}$. Since $1 \in E$, E is non-empty. Since $n > 0$, E is bounded above by na. Hence, $k_0 := \sup E$ exists and belongs to E, in particular to $\mathbb{N}$.

Set $m = k_0 + 1$ and $q = \frac{m}{n}$. Since k_0 is a supremum of E, m is not in the domain of E. Thus, $q > a$.

On the other hand, since k_0 is in the domain of E, it follows from our choice of n that

$$b = a + (b - a) \geq \frac{k_0}{n} + (b - a) > \frac{k_0}{n} + \frac{1}{n} = \frac{m}{n} = q.$$

Now suppose that $a \leq 0$, choose, by the archimedean principle $k \in \mathbb{N}$ such that $k > -a$ then $0 < k + a < k + b$, and by the case already proved, there is an $r \in \mathbb{Q}$ such that $k + a < r < k + b$. Therefore, $q := r - k$ belongs to $\mathbb{Q}$ and satisfies the inequality $a < q < b$. $\square$

What follows is a quick theorem regarding a relationship between suprema and infima.

Theorem 8.11. *Let $E \subset \mathbb{R}$ be non-empty.*

- *E has a supremum if and only if $-E$ has an infimum, in which case*

$$\inf(-E) = -\sup E$$

- *Likewise, E has an infimum if and only if $-E$ has a supremum, in which case*

$$\sup(-E) = -\inf E.$$

For our last theorem of this section, we have chosen a broadly applicable theorem which will be very useful later on.

Theorem 8.12. *Suppose that $A \subset B$ are non-empty subsets or $\mathbb{R}$.*

1. *If B has a supremum, then* $\sup A \leq \sup B$.
2. *If B has an infimum, then* $\inf A \geq \inf B$.

Proof. i) Since $A \subset B$, any upper bound of B is an upper bound of A. Therefore, $\sup B$ is an upper bound of A. It follows from the completeness axiom that $\sup A$ exists, and by necessity, it cannot exceed the supremum of B.
ii) Clearly, $-A \subset -B$. Thus, by our previously proved statement

$$\inf A = -\sup(-A) \geq -\sup(-B) = \inf B.$$

$\square$

8.2 Sequences

An *sequence* is a function whose domain is $\mathbb{N}$. A sequence f whose terms are $x_n := f(n)$ will be referenced by a sequence of terms $(x_1, x_2, ...)$ or by $\{x_n\}$. Thus 1, 1/2, 1/4, 1/8 , ... is a representation of the sequence $\{1/2^{n-1}\}_{n \in \mathbb{N}}$ and 1, 2, 3, 4, ... is a representation of the sequence $\{n\}_{n \in \mathbb{N}}$.

We can make an effort to bound sequences in much the way same way as we would bound sets. In fact, we do so by looking at the bounds of the infinite sets that the sequences form.

Definition 8.13. *Let $\{x_n\}$ be a sequence of real numbers*

i) The sequence $\{x_n\}$ is said to be bounded above *if and only if the set$\{x_n : n \in \mathbb{N}\}$ is bounded above.*

ii) The sequence $\{x_n\}$ is said to be bounded below *if and only if the set $\{x_n : n \in \mathbb{N}\}$ is bounded below.*

iii) $\{x_n\}$ is said to be bounded *if and ony if it is bounded above and bounded below.*

Without getting too bogged down too heavily in the idea of convergence, let us state a quick theorem regarding the relationship between bounding and convergence.

Theorem 8.14. *Every convergent sequence is bounded.*

Proof. Given $\varepsilon = 1$, there is an $N \in \mathbb{N}$ such that $n \geq N$ implies $|x_n - a| < 1$. Hence by the triangle inequality, $|x_n| < 1 + |a|$ for all $n \geq N$. On the other hand, if $1 \leq n \leq N$, then.

$$|x_n| \leq M := max\{|x_1|, |x_2|, ..., |x_n|\}$$

Therefore, $\{x_n\}$ is dominated by $max\{|x_1|, |x_2|, ..., |x_n|\}$ □

The limit concept is one which is absolutely fundamental to analysis. As such, we would do well to have a functional definition of the limit of a sequence.

Recall that in basic calculus we say that a sequence of real numbers $\{x)n\}$ converges to a number a if x_n gets arbitratily near a (i.e. the distance between a and x_n gets small) as n gets large. Thus, given $\varepsilon > 0$ (no matter how small), if n is large enough, $|x_n - a|$ is smaller than ε. This leads us to the following formalized definition of the limit of a sequence.

Definition 8.15. *A sequence of real numbers $\{x_n\}$ is said to* converge *to a real number $a \in \mathbb{R}$ if and only if $\forall \varepsilon > 0 \exists N \in \mathbb{N}$ (which, in general, depends upon ε) such that*

$$n \geq N \Rightarrow |x_n - a| < \varepsilon$$

8.2.1 Subsequences

We may sometimes find it useful to cut a sequence into bits. These subsequences can be useful in a variety of ways. A subsequence may be defined as follows:

Definition 8.16. *By a* subsequence *of a sequence $\{x_n\}_{n\in\mathbb{N}}$, we shall mean a sequence of the form $\{x_{n_k}\}_{n\in\mathbb{N}}$, where $n_k \in \mathbb{N}$ and $n_1 < n_2 < ...$*

Thus a subsequence is obtained by removing values from the sequence. For example, we could take the sequence $\{-1^n\}$ and obtain the subsequences $-1, -1, -1, -1, ...$ and $1, 1, 1, 1, ...$

8.3 Limits of Functions and Continuity

Of central concern to the study of Analysis is whether a given sequence converges. We will go into several theorems that aid us in deciding on the status of a sequence.

Theorem 8.17. Squeeze Theorem
Suppose that $\{x_n\}$, $\{y_n\}$, and $\{w_n\}$, are real sequences.

i) If $x_n \to a$ and $y_n \to a$ (the same a) as $n \to \infty$, and if there is an $N_0 \in \mathbb{N}$ such that

$$x_n \leq w_n \leq y_n, n \geq N_0$$

then $w_n \to a$ as $n \to \infty$
ii) If $x_n \to 0$ as $n \to \infty$ and $\{y_n\}$ is bounded, then $x_n y_n \to 0$ as $n \to \infty$

Let us go very quickly through a proof regarding this, since it is one of our most commonly used theorems.

Proof. part i) Let $\varepsilon > 0$ since x_n and y_n converge to a, choose $N_1, N_2 \in \mathbb{N}$ such that, $n \geq N_1$ implies $-\varepsilon < x_n - a < \varepsilon$ and $n \geq N_2$ implies $-\varepsilon < y_n - a < \varepsilon$. Set $N = max\{N_0, N_1, N_2\}$. If $n \geq N$ we have by hypothesis and by the choice of N_1 and N_2 that

$$a - \varepsilon < x_n \leq w_n \leq y_n < a + \varepsilon;$$

that is, $|w_n - a| < \varepsilon$ for $n \geq N$. We conclude that $w_n \to a$ as $n \to \infty$.

part ii) Suppose that $x_n \to 0$ and that there is an $M > 0$ such that $|y_n| \leq M$ for $n \in \mathbb{N}$. Let $\varepsilon > 0$ and choose an $N \in \mathbb{N}$ such that $n \geq N$ implies $|x_n| < \varepsilon/M$. Then $n \geq N$ implies

$$|x_n y_n| < M\frac{\varepsilon}{M} = \varepsilon.$$

We conclude that $x_n y_n \to 0$ as $n \to \infty$. □

Let us look at an example where the squeeze theorem makes our life extremely easy.

Example 8.18. *Find $lim_{n\to\infty} 2^{-n}\cos(n^3 - n^2 + n - 13)$*

Solution. *The factor $\cos(n^3 - n^2 + n - 13)$ looks intimidating, but it is superfluous when we are looking for the limit of this sequence.*
Indeed, since $|\cos(x)| \leq 1$ for all $x \in \mathbb{N}$, the sequence $2^{-n}\cos(n^3 - n^2 + n - 13)$ is dominated by the 2^{-n} term.
Since $2^n > n$, it is clear by the squeeze theorem that both $2^{-n} \to 0$ and that $2^{-n}\cos(n^3 - n^2 + n - 13) \to 0$ as $n \to \infty$

Here is an extremely handy theorem which focuses on breaking up sequences in order to make finding limits more manageable:

Theorem 8.19. *Suppose that $\{x_n\}$ and $\{y_n\}$ are real sequences and that $\alpha \in \mathbb{R}$. If $\{x_n\}$ and $\{y_n\}$ are convergent, then*

1. $\lim_{n\to\infty}(x_n + y_n) = \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n$
2. $\lim_{n\to\infty}(\alpha x_n) = \alpha \lim_{n\to\infty} x_n$
3. $\lim_{n\to\infty}(x_n y_n) = (\lim_{n\to\infty} x_n)(\lim_{n\to\infty} y_n)$
   If, in addition, $y_n \neq 0$ and $\lim_{n\to\infty} y_n \neq 0$, then
4. $\lim_{n\to\infty} \frac{x_n}{y_n} = \frac{\lim_{n\to\infty} x_n}{\lim_{n\to\infty} y_n}$

(in particular, all of these limits exist.)

These theorems are quite handy for breaking up a sequence in order to decide whether it converges or not. On the subject of a sequence not converging, it would be well for us to have an adequate definition of divergence going forward.

Definition 8.20. *Let $\{x_n\}$ be a sequence of real numbers*

1. *$\{x_n\}$ is said to* diverge *to $+\infty$ (notation: $x_n \to +\infty$ as $n \to \infty$ or $\lim_{n\to\infty} x_n = +\infty$ if and only if for each $M \in \mathbb{R}$ there exists an $N \in \mathbb{N}$ such that*
$$n \geq N \to x_n > M.$$
2. *$\{x_n\}$ is said to* diverge *to $-\infty$ (notation: $x_n \to -\infty$ as $n \to \infty$ or $\lim_{n\to\infty} x_n = -\infty$ if and only if for each $M \in \mathbb{R}$ there exists an $N \in \mathbb{N}$ such that*
$$n \geq N \to x_n < M.$$

Notice that by this definition $x_n \to +\infty$ if and only if given $M \in \mathbb{R}$, x_n is greater than M for a sufficiently large n; that is that x_n eventually exceeds M, no matter how arbitrarily large. Likewise, $x_n \to -\infty$ if and only if x_n is eventually less than M, no matter how large and negative it may be.

Now we will state a theorem regarding divergent sequences which comes in very handy when dealing with sequences which contain divergent pieces.

Theorem 8.21. *Suppose that $\{x_n\}$ and $\{y_n\}$ are real sequences such that $x_n \to +\infty$ (respectively, $x_n \to -\infty$) as $n \to \infty$.*

1. *If y_n is bounded below (respectively, y_n is bounded above), then*
$$\lim_{n\to\infty}(x_n + y_n) = +\infty \text{ (respectively, } \lim_{n\to\infty}(x_n + y_n) = -\infty)$$
2. *If $\alpha > 0$, then*
$$\lim_{n\to\infty}(\alpha x_n) = +\infty \text{ (respectively, } \lim_{n\to\infty}(\alpha x_n) = -\infty)$$

3. If $y_n > M_0$ for some $M_0 > 0$ and all $N \in \mathbb{N}$, then

$$\lim_{n\to\infty}(x_n y_n) = +\infty \text{ (respectively, } \lim_{n\to\infty}(x_n y_n) = -\infty)$$

4. If $\{y_n\}$ is bounded and $x_n \neq 0$, then

$$\lim_{n\to\infty} \frac{y_n}{x_n} = 0.$$

Proof. We suppose for simplicity that $x_n \to +\infty$ as $n \to \infty$

1. By hypothesis, $y_n \geq M_0$ for some $M_0 \in \mathbb{R}$. Let $M \in \mathbb{R}$, set $M_1 = M - M_0$. Since $x_n \to +\infty$ choose $N \in \mathbb{N}$ such that $n \geq N$ implies $x_n > M_1$. Then $n \geq N$ implies $x_n + y_n > M_1 + M_0 = M$.

2. Let $M \in \mathbb{R}$ and set $M_1 = \frac{M}{\alpha}$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies $x_n > M_1$. Since $\alpha < 0$ we concluded that $\alpha x_n > \alpha M_1 = M \; \forall n \geq N$

3. Let $M \in \mathbb{R}$ and set $M_1 = \frac{M}{M_0}$ Choose $N \in \mathbb{N}$ such that $n \geq N$ implies that $x_n > M_1$. Then $n \geq N$ implies that $x_n y_n > M_1 M_0 = M$.

4. Let $\varepsilon > 0$. Choose $M_0 > 0$ such that $|y_n| \leq M_0$ and $M_1 > 0$ sufficiently large that $\frac{M_0}{M_1} < \varepsilon$. Choose $N \in \mathbb{R}$ such that $n \geq N$ implies $x_n > M_1$. Then $n \geq N$ implies that

$$|\frac{y_n}{x_n}| = \frac{|y_n|}{x_n} < \frac{M_0}{M_1} < \varepsilon.$$

$\square$

If we adopt the conventions

$$\begin{array}{ccc} x + \infty = \infty & x - \infty = -\infty & x \in \mathbb{R} \\ x * \infty = \infty & x * -\infty = -\infty & x > 0 \\ x * \infty = -\infty & x * -\infty = \infty & x < 0 \\ \infty + \infty = \infty & -\infty - \infty = -\infty & \\ \infty * \infty = & -\infty * -\infty = & \infty \\ \infty * -\infty = & -\infty * \infty = & -\infty \end{array}$$

Then our theorem contains the following corollary.

Corollary 8.22. *Let $\{x_n\}, \{y_n\}$ be real sequences and α, x, y be extended real numbers. If $x_n \to x$ and $y_n \to y$ as $n \to \infty$ then*

$$\lim_{n\to\infty} x_n y_n = x + y$$

Provided that the right side is not of the form $\infty - \infty$, and

$$\lim_{n\to\infty} \alpha x_n = \alpha x, \lim_{n\to\infty} x_n y_n = xy$$

Provided that none of these products is of the form $0 * \pm\infty$

We see from these theorems that a divergent piece or term in a sequence is enough to make the entire sequence divergent (or converge to 0 if it happens to be a denominator). In effect, it can poison an otherwise well behaved sequence, and as such is something to look out for.

Last but not least, we will conclude our discussion of Real Analysis with the Comparison Theorem, which is extremely useful for bounding sequences.

Theorem 8.23. *Suppose that $\{x_n\}$ and $\{y_n\}$ are convergent sequences. If there is an $N_0 \in \mathbb{N}$ such that*

$$x_n \leq y_n, \forall n \geq N_0$$

then

$$\lim_{n\to\infty} x_n \leq a \lim_{n\to\infty} y_n$$

In particular, if $x_n \in [a, b]$ converges to some point c then c must belong to $[a, b]$.

Proof. Suppose that the first statement is false; that is to say that

$$x_n \leq y_n, \forall n \geq N_0$$

holds but $x := \lim_{n\to\infty} x_n$ is greater than $y := lim_{n\to\infty} y_n$. Set $\varepsilon = \frac{x-y}{2}$. Choose $N_1 > N_0$ such that $|x_n - x| < \varepsilon$ and $|y_n - y < \varepsilon$ for $n \geq N_1$. Then for such an n,

$$x_n > x - \varepsilon = x - (\frac{x-y}{2}) = y + (\frac{x-y}{2}) = y + \varepsilon > y_n,$$

which contradicts the statement we made in the beginning ($x_n \leq y_n, \forall n \geq N_0$). This proves our first statement.

We conclude by noting that the second statement follows from the first, since $a \leq x_n \leq b$ implies that $a \leq c \leq b$ ∎

One way to remember this result is to say that the limit of an inequality is the inequality of the limits, provided that these limits exist. We shall call this process taking the limit of an inequality. The following corollary contains an implication of the comparison theorem which is important to remember.

Corollary 8.24.

$$x_n < y_n \to \lim_{n\to\infty} x_n \leq \lim_{n\to\infty} y_n$$

However, this does NOT mean the following

$$x_n < y_n \to \lim_{n\to\infty} x_n < \lim_{n\to\infty} y_n.$$

We do not necessarily know what x and y converge to, so this statement does not work.

9 Statistics

Statistics has always had an interesting relationship with mathematics. In some schools, statistics departments are held separate from the math department. Regardless, statistics is another valuable tool in a mathematicians arsenal.

The information for the following summary of elementary statistics is taken from [6]. Please note that the examples are also drawn from this text as I no longer have my notes from the course.

9.1 Introduction to statistics

Statistics is the branch of mathematics concerned with the collecting, organizing, analysing, and interpreting of data in order to gain some insight regarding a population or make decisions.

This is an extremely broad definition, and we would do well to firmly define a couple of the terms in it.

Definition 9.1. *Data: Consists of information coming from observations, measurements or responses.*

From this we see that data can take many forms, both social and physical. Next we must firmly define what populations and samples are.

Definition 9.2. *Population: The collection of* all *outcomes, responses, measurements or counts that are of interest. Examples include: All people in a country, the set of all cars in the US, and the student body of a college.*

Definition 9.3. *Sample: a subset or part of the population.*

Note that our definition of sample gives no warranty as to the capacity of a sample to be representative of its population. In fact, one of the earliest problems covered in statistics is how to take a good sample that is representative of the whole. Without a good sample, everything else we do in statistics is effectively worthless.

There are two more important definitions that we will often use in our discussions of statistics, namely parameters and statistics, which can be defined as follows:

Definition 9.4. *Parameter: A numerical description of a population characteristic.*

Definition 9.5. *Statistic: A numerical description of a sample characteristic.*

If we have a good sample (i.e. one that as representative of the population) then our statistic will correspond strongly with a parameter regarding our population. As such, we can determine some characteristic about even an exceedingly large population with some degree of confidence based on a much more manageable smaller sample.

Now that we have defined the basic terminology of statistics, let us examine the two major branches of statistics.

Descriptive statistics is the branch of statistics which involves the organization, summarization and displaying of data.

Inferential statistics is the branch of statistics concerned with using a sample to draw conclusions regarding a population. A basic tool of inferential statistics is probability, which we may discuss later.

9.1.1 Data

Lets get more granular in our discussion of data. Data can be broadly defined as being one of two types.

Definition 9.6. *Qualitative Data: Data which consists of attributes, labels, or non-numerical features. Things like hair color, race, or political affiliation would be of this type.*

Definition 9.7. *Quantitative Data: Data which consists of numerical measurements or counts. Features such as age, GMAT scores, or gross income would be of this type.*

As a rule of thumb, if it makes sense to try and find the mean, you're probably dealing with quantitative data. Here is a finer division of types of data.

Theorem 9.8. *Levels of measurement*

1. *Nominal level: Data at the nominal level of measurement is qualitative only. Data may be categorized using names, labels or qualities. No mathematical computations may be made at this level.*

2. *Ordinal Level: Data at the ordinal level of measurement may be qualitative or quantitative. Data at this level may be arranged, ordered or ranked, but differences between data entries are not meaningful. Examples include SSID numbers.*

3. *Interval Level: Data at the interval level of measurement may be ordered, and has meaningful differences between data entries which can be calculated. At the interval level, a zero entry simply represents a position on scale; the entry is not an inherent zero.*

4. *Ratio Level: Data at the ratio level of measurement is broadly similar to data at the interval level with the added property that a zero entry is an inherent zero. A ratio of two or more entries can be formed so that one entry may be meaningfully expressed as a multiple of another.*

Now we will move on to one of the most important topics in statistics: Data collection.

9.1.2 Data Collection

Data collection can be done in several ways. Often, the type of data collection used is dictated by the type of study being done.

A *simulation* is the use of a mathematical or physical model to reproduce the conditions of a situation or process. Collecting data often involves the use of computers. Simulations allow you to study situations that are impractical or even dangerous to create in real life, and often save time and money. For instance, automobile manufacturers use simulations with dummies to study the effect of crashes on humans.

A *Survey* is an investigation of one or more characteristics of a population. Most often, surveys are carried out on people by asking them questions. The most common types of surveys are done by interviews, internet, phone or mail. In designing a survey, one must be careful to write the questions in such a way that they do not lead to a biased results. For example, when surveying to determine attitudes towards a rezoning ordinance, one shouldn't ask a leading question such as: do you feel that our community would be served by more liquor stores near our children's schools?

When one does a survey or a simulation, they must be aware of the effect of a confounding variable upon their study.

Definition 9.9. *Confounding variable: A confounding variable occurs when an experimenter cannot tell the difference between the effects of different factors on the variable.*

As an example, if a coffee shop remodels to increase the flow of business and a shopping mall opens next door leading to increased foot traffic, it would be difficult to determine what was the procuring cause of any increased business.

A common cause for error in the medical industry is the idea of the placebo effect.

Definition 9.10. *Placebo effect: occurs when a subject reacts favourably to a placebo when in fact they have been given a fake treatment. This can make it difficult to determine the efficacy of the actual treatment being tested.*

In order to help control and minimize the placebo effect, we can put blinding procedures in place.

Definition 9.11. *Blinding: A technique in which some party is not aware whether they are receiving a treatment or a placebo. To take this a step further and remove any sort of bias, trials may be double blind, meaning that the doctor administering the treatment does not know whether the subject is getting a placebo or a genuine treatment.*

Now let us move on to the realm of descriptive statistics.

9.2 Descriptive Statistics

There are several things to look at when examining a set of data. Some important characteristics include its center (where do data points tend to gravitate towards), its variability (how far do data points spread away from our center), and its shape (normal, bimodal, etc). This brings us to our first statistics.

9.2.1 Mean, Median, and Mode

A measure of central tendency is a value that represents the typical entry in a set of data. This is a figure we would very much like to know about, and as such, we have 3 very basic methods of determining the typical point in a data set.

A *mean* is the sum of a group of data points divided by the number of data points in the set. Or, to put it in mathematical terminology

$$\frac{\Sigma x}{n}$$

where n is the number of data points in the set being averaged.

The *median* of a data set is the point which lies in the middle of a set when it is ordered numerically. If the data set has an even number of elements, and no point lies precisely in the middle, then the median is the average of the two middle-most data points.

The *mode* of a data set is the element(s) which occur with the greatest frequency. Note that a set may have 0,1,2 or more modes.

In beginning our discussion of descriptive statistics, we come now to the hairy subject of outliers.

Definition 9.12. *An* outlier *is a data entry that is far removed from the other entries in a data set. We may later discuss statistically valid ways to remove outliers from our data sets.*

We may also do what is called a weighted mean, where each entry is given a comparative weight which is factored into the final mean score. This looks like

$$\frac{\Sigma x * w}{n}.$$

An example of a weighted mean that is familiar to all students is the varying weights given to various components of coursework. For example, our final exam may be worth 25% of our grade, while homework may count for only 10%.

Now we move on to two extremely fundamental concepts in the realm of descriptive statistics: standard deviation and variance., both of which are measures of the average data point's deviance from the mean.

The *Population Variance* is defined as follows

$$\sigma^2 = \frac{\Sigma(x-\mu)^2}{N}$$

However, there is one distinct disadvantage to variance: it does not have the same units as the thing measured; in fact its units are squared. To get around this issue, we take the square root, giving us the *Population Standard Deviation*

$$\sigma = \sqrt{\frac{\Sigma(x-\mu)^2}{N}}.$$

Please note that these formulas work identically for a sample instead of a population.

There is a quick rule of thumb regarding standard deviations and the normal curve which bears mentioning.

Theorem 9.13. The 68-95-99.7 Rule

1. *Approximately* 68% *of the data lies within one standard deviation of the mean.*
2. *Approximately* 95% *of the data lies within two standard deviations of the mean.*
3. *Approximately* 99.7% *of the data lies within three standard deviations of the mean.*

Please note that this rule only applies to data which has a roughly normal distribution. This may seem quite esoteric, but a shocking number of things in the real world follow a normal curve with regards to distribution.

Last but not least, we will mention the coefficient of variation. The coefficient of variation is a nifty statistic which expresses standard deviation as a percentage of the mean. In a population, it would look like so

$$\frac{\sigma}{\mu}.$$

References

[1] Mactutor history of mathematics. http://www-history.mcs.st-and.ac.uk/. Accessed: 2015-03-15.

[2] Steven L Campbell and Richard Haberman. *Introduction to Differential Equations with Dynamical Systems*. Princeton University Press, 2008.

[3] Peter J Eccles. *An Introduction to Mathematical Reasoning*. Cambridge University Press, 1st edition, 2007.

[4] Frank Giordano, Maurice Weir, and William Fox. *A First Course in Mathematical Modelling*. 3rd edition.

[5] R. Johnson and F. Kiokemeister. *Calculus with Analytic Geometry*. ALLYN and BACON, 1960.

[6] Ron Larsen and Betsy Farber. *Elementary Statistics, Picturing the World*. Pearson, 6th edition, 2015.

[7] Seymour Lipschutz and Marc Lipson. *Schaum's Outline of Theory and Problems of Linear Algebra*. McGraw Hill, 3rd edition, 2001.

[8] Neal McCoy and Gerald Janusz. *An Introduction to Abstract Algebra*. Trustworthy Communications, 7th edition, 2009.

[9] J. Saxon and F. Wang. *Calculus with Trigonometry and Analytic Geometry*. Saxon Publishers, 1997.

[10] James Stewart. *Calculus Early Transcendentals*. 2008.

[11] William R. Wade. *An Introduction to Analysis*. Prentice Hall, 4th edition, 2010.