IUP Applied Mathematics and Data Science

A Survival Guide for
Applied Mathematics and Data Science

Updated: January 27, 2022

Edited by John Chrispell

Contributing Authors:
Kimberly Burch, Shelly Bouchat, Yongtao Cao,
John Chrispell, Alfy Dahma, Tim Flowers,
Yu-Ju Kuo, Russ Stocker, and Gary Stoudt

Department of Mathematical and Computer Sciences
Kopchick College of Natural Sciences and Mathematics
Indiana University of Pennsylvania

Preface

The contents in the document are a crash course in the material that a student will need to be successful when starting out as a graduate student in Applied Mathematics and Data Science. The contents contained herein are by no means all encompassing. They only serve as a refresher on some topics that for some students may have started to fade. These materials will serve as starting exposure to topics that will help students finding their path when starting graduate studies.

Table of Contents

Chapter 1

Mathematical Typesetting

LaTeX is a a markup language commonly used in mathematical typesetting. The typsetting system as it is currently known started with Donald Knuth in 1977 when he started writting TeX. The idea was to create a powerful and flexible typsetting utility. The markup system that Knuth created is especially good for mathematics.

Modern LaTeXis the product of Leslie Lamport (1980) where commands were added over the standard TeXThe for LaTeXwas to separate content from style enabling structured documents. The power of this form of document typesetting is that it will automates numbering, cross-referencing, bibliography, index, etc, so that the author doesn't need to continually and manually track and update these items.

The advantages of using LaTeX for mathematics are many:

1. The tool is free!

2. Designed for scientific texts!

3. It is a structured document markup language.

4. Gives authors much greater control of formatting when compared to other document formatting systems.

5. The tool is easily scalable.

6. Users can program in the markup language and thus easily extend it.

7. Documents created in LaTeX are portable.

1.1　Software Packages and Setting Up LaTeX Editors

To get started without the pain of installing software, their are several web based programs where you can typeset in LaTeX Overleaf is a common place for new users to start that offers templates that can be adapted to most situations. New users are recommended to check out:

`https://www.overleaf.com`

as a key starting place.

Users can also typeset LaTeX documents directly on CoCalc.

`https://cocalc.com/`

LaTeX on Windows is typically consists of installing the collection of classfiles defined in a MikTeXdistribution. Specifically MikTeXis a set of macros for TeX. It includes LaTeX and other packages needed for document typesetting. This collection can be downloaded from:

`www.miktex.org`

TeXLive is also an easy way to get up and running with the TeX document production system.

`www.tug.org/texlive/`

1.1.1　Front-End Editors

As LaTeX is a typsetting language the editor used to create documents is not important, and documents can be written on any editor that allows for organizing text. Some editors have been created with editing LaTeX documents in mind. A list of LaTeX editors is available at:

`beebom.com/best-latex-editors/`

A common editor used on multiple platforms is TexMaker.

`www.xm1math.net/texmaker/index.html`

A common Windows editor is WinEdt `www.winedt.com`

Mac users sometimes enjoy TexShop as their editor.

There are many other text editors for LaTeX for example LyX. Some are embedded in mathematical software, such as Mathematica, Matlab, and scientific notebook.

1.2 Typesetting a Document

All documents need the following header commands:

Example 1.1: Basic Document

```
\documentclass[12pt]{article}
\begin{document}
The text of your document.
\end{document}
```

To create sections and subsections use:

Example 1.2: Sections

Use commands `\subsection{A subsection}` Note subsection without an asterisk. `\subsection*{A subsection too}` This will not be in the table of contents.

In (1.2) we do some stuff. In (1.1) we do other stuff. Citing things in the document we use the `\cite{dut961}` in the text [**?**] to get the citation and add it to the bibliography.

The Package works as a markup language for the text you type. It has two types of equation environment to use. The first is the in paragraph math environment. To use this environment place dollar signs around your desired mathematical expression.

Example 1.3: Inline Math

For example putting `$\frac{stuff}{otherstuff}$` in a line of text would produce: $\frac{stuff}{otherstuff}$ in the line of text.

The second type of expression environment lets you place mathematics expressions on separate lines. Use `\[ mathematical expression \]` to do this.

Example 1.4: Separate Line of Math

The following line would produce the expression that follows.

```
\[ \int_0^1{\frac{1}{1+x^2}}dx = \frac{\pi}{4} \]
```

$$\int_0^1 \frac{1}{1+x^2} dx = \frac{\pi}{4}$$

Other key items: Things are what you think they should be! `\sum` produces $\sum$. Typing `\underline{ stuff }` will produce the result: <u>stuff</u>. The braces are used by LaTeX as a form of unseen parentheses.

1.3   Inside the math environment:

1. Spacing in the math environment: Look at the term

$$\int_0^1 \frac{1}{1+x^2} dx = \frac{\pi}{4}$$

   this looks much better written as:

$$\int_0^1 \frac{1}{1+x^2} \, dx = \frac{\pi}{4}$$

   To achieve this we use `\!` which makes the space between items in the math environment smaller, and `\;` which make the space between items greater.

2. Typing text in an equation. Use the command `\mbox{ text }` to place text such as 'and' in the expression $x \leq 2$ and $x \geq 25$.

1.4   Spacing

1. To move to various distances in a line use the `\vskip1in` command or use the `\hskip1.2in` to leave a blank space of 1.2 inches. LaTeX recognizes the spacing commands in either inches or in points (pt).

2. To double space or one and a half space a document use the command `\baselineskip24pt` or whatever distance is desired.

3. To play with the margins use the command `\oddsidemargin1in` . The top margin is `\topmargin1in`, and the bottom margin is adjusted using the term `\textheight8in` . Adjust the right hand margin using `\textwidth6in`. Note: LaTeX will recognize negative spacing in these commands.

4. The command `\\` will end a line and start a new one.

5. To start a new paragraph just simply skip a line in you text document.

6. LATEX will automatically indent the first line of a new paragraph. To stop this use the command `\noindent`.

7. To get rid of pesky page numbers centered at the bottom of the page use the command `\thispagestyle{empty}` .

1.5 Changing Text style

1. The command `\tt` will produce `type writer style text`.

2. The command `\bf` will produce the **bold face text.**, and `\em` will produce *nice typing* .

3. The size of text can be changed using the slash and the following words: tiny, scriptsize, footnotesize, small, normalsize, large, Large, LARGE, huge, and HUGE. Note: These changes in the text size stay until you change them back.

1.6 Some Basic Environments

1. The center environment: Use the command `\begin{center}` then type your information, and when your done just close the environment using the command `\end{center}` . Note: All environments must be closed otherwise LATEX will give an error; however, it is nice and shows where it thinks the Error occurred.

2. The Quote Environment: Works the same as the center environment. Place your quote between `\begin{quote}` and `\end{quote}`.

3. The Enumerate Environment: This works as a way to number items, such as this list. Begin the environment using `\begin{enumerate}` , then to enter new items in the list use `\item` . Note: Nested enumerates may be used.

4. The array environment: This will allow you to set equations equal, and to type matrices, and tables. It is best explained by just looking at an example and comparing with the output.

$$\left(\begin{array}{c|c c} 4x^2 & 3y^3 & = \left| \begin{array}{c} 0 \\ . \\ \dots \end{array} \right. \\ x_1 & \vdots & = \\ f_{n+3}(x) & f_{n+2} & = \end{array} \right]$$

The LATEX markup to get this expression is:

```
\[ \left( \begin{array}{c|rc|l}         \\
4x^2       &   3y^3   & = & 0        \\
x_1        &   \vdots  & = & \cdot  \\
f_{n+3}(x) &   f_{n+2} &  = & \ldots \\
\end{array} \right] \]
```

Note: that the & separates the boxes, within the array. Also the r, c, and l give the alignment of each column. The \\ term is used to end the row.

A second Example:

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

To get the above expression just type the following:

```
\[ \left[ \begin{array}{cc}
   2  &  4 \\
   3  &  1 \\ \end{array}\right]
+ \left[ \begin{array}{cc}
   1  &  0  \\
   0  &  1  \\ \end{array}\right]
= \left[ \begin{array}{cc}
   3  &  4 \\
   3  &  2 \\ \end{array} \right] \]
```

1.7   Figures

LATEX allows for .eps files to be imported nicely into a figure environment. One only needs to export the figure as file.eps form its source. The picture below can be achieved using the following text:

```
\begin{figure}[htp]
\centering
   \subfigure[logo 1]{
        \label{p-s-plot_sand}
        \includegraphics[angle=0, height=1.0in]{iuplogo}}
   \hspace{.75in}
   \subfigure[logo 2]{
        \label{p-s-plot_all}
        \includegraphics[angle=90, height=1.0in]{iuplogo}}
   \caption{Two logos! \label{logoB}}
\end{figure}
```

(a) logo 1

(b) logo 2

Figure 1.1: Two logos!

Note: you must have :
```
  \usepackage[dvips]{graphics}
 \usepackage[dvips]{graphicx}
```

in the preamble of your document for the above code to work.

1.8 Including TEX from Other files

There are two ways to do this nicely:

1. Use the `\include{filename}` command. This option will include all the TEX in the given file and also inserts a page break before the markup is included.

2. Use the `\input{filename}` command. This brings your markup in without any altered spacing commands.

Note in both cases you need not include any of the preamble.

```
\begin{document} ...
\end{document}
```

1.9 Cross Referencing Expressions or Sections

In order to reference an equation or expression you can use a labeling command. The following

```
\begin{equation}
\nabla {\bf u} \cdot  {\bf g} = {\bf 0}.
\label{refeqq}
\end{equation}
```

will produce

$$\nabla\mathbf{u}\cdot\mathbf{g}=\mathbf{0}. \qquad (1.1)$$

Note that the expression has a reference in it. You can then reference this later in the text by using a reference to the label with `\ref{refeqq}` as in expression (1.1) some stuff is done. Using an asterisk will suppress the equation numbering.

```
\begin{equation*}
\Delta {\bf u} \cdot  {\bf g} = {\bf 0}
\end{equation*}
```

Note that

$$\Delta\mathbf{u}\cdot\mathbf{g}=\mathbf{0}$$

the equation directly above has no label. A similar effect can be accomplished using the `\nonumber` command. The markup given by:

```
\begin{eqnarray}
 y & = & \pi x + 2 \nonumber \\
    & = & 7 \Pi \label{refeqb}
\end{eqnarray}
```

produces the typeset expression:

$$\begin{aligned} y &=& \pi x + 2 \\ &=& 7\Pi \end{aligned} \qquad (1.2)$$

1.10   Slides and Presentations

The "beamer" package is quite handy for making your LaTeX markup documents into slides for talks and presentations. After ensuring that the beamer class for LaTeX is installed you can then typeset slides as follows:

```
\documentclass{beamer}
\usepackage{beamerthemeshadow}
\begin{document}
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
\begin{frame}
\frametitle{\LaTeX Commands}
Generally speaking, there are three types of commands:
\begin{itemize}
\item Simple commands: start with a $\backslash$ followed directly
        by the command\\
    ex: $\backslash$newpage
```

```
    \pause
\item Environments: start with a $\backslash$begin$\{...\}$ and
          end with $\backslash$end$\{...\}$\\
    ex: $\backslash$begin$\{$equation$\}$\\
    $\backslash$frac$\{$dy$\wedge$2$\}$$\{$d$\wedge$2x$\}$
    +$\backslash$frac$\{$dy$\}$$\{$dx$\}$=x$\wedge$2+1\\
    $\backslash$end$\{$equation$\}$
    \begin{equation}
    \frac{dy^2}{d^2x}+\frac{dy}{dx}=x^2+1
    \end{equation}\pause

\item Commands that require additional arguments\\
ex: $\backslash$textbf$\{$ABC$\}$\rightarrow$ \textbf{ABC}

\end{itemize}
\end{frame}
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
\end{document}
```
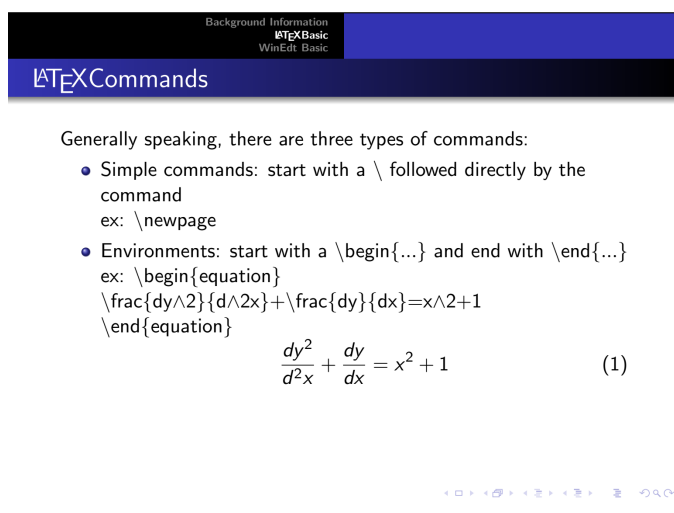
The above given markup produces a slide that is shown in figure 1.2. Note the following:

- The document class is defined to be beamer.

- You can easily change the theme of the slides by adjusting theme used.

- Each slide in your presentation is defined using the "frame" command.

- Pauses or breaks can be included in the slide by using the "\pause" command.

The reader is advised to do some searching on the beamer package and how it can be used.

Figure 1.2: Example of a beamer slide.

Chapter 2

Calculus Review

This chapter will serve as a brief review of the essential concepts encountered in a calculus sequence.

2.1 Sequences

Assuming that the reader is familiar with functions and basic function notation our review will start with sequences and the set of real number $\mathbb{R}$ and a discussion of limits.

A sequence in $\mathbb{R}$ is usually thought of as a list of numbers:

$$x_1, x_2, x_3, \ldots, x_n, x_{n+1}, \ldots$$

where $x_i \in \mathbb{R}$ for all i in $\mathbb{N}$.

Example 2.1: Basice Sequence

$$(x_n)_{n \in \mathbb{N}} = \left(\frac{1}{n}\right)_{n \in \mathbb{N}} = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \frac{1}{n+1}, \ldots\right)$$

It is often useful to view a sequence as a function that maps $\mathbb{N}$ to the real numbers $\mathbb{R}$. That is:

Definition 2.1: Sequence

A sequence is a function $f : \mathbb{N} \to \mathbb{R}$.

In the above Example 2.1 the list of numbers is then the range of the sequence $f : \mathbb{N} \to \mathbb{R}$ defined by $f(n) = \frac{1}{n}$, $n \in \mathbb{N}$. Here we see that

$$f(1) = x_1 = 1, f(2) = x_2 = \frac{1}{2}, \ldots$$

> **Definition 2.2: Limit**
>
> A real number $x \in \mathbb{R}$ is the <u>limit</u> of the sequence $(x_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ if for every $\epsilon > 0, \exists\, N \in \mathbb{N}$ such that $|x_n - x| < \epsilon,\ \forall\, n \geq N$.

Using the Definition 2.2 we say that (x_n) converges to x and we can write:

$$\lim_{n \to \infty} (x_n) = x \text{ or } x_n \to x, n \to \infty.$$

To expand on the definition of a limit consider a fixed value of $\epsilon > 0$ where ϵ is set arbitrarily as small as you desire, then visually we see the convergent sequence as:
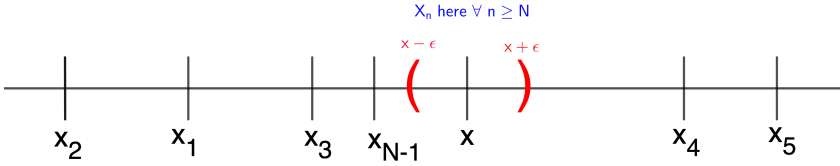
In Example 2.1 $x_n = \frac{1}{n} \to 0$ as $n \to \infty$. Indeed, fix any $\epsilon > 0$ as small as you like. Let $N \in \mathbb{N}$ be defined such that $N > \frac{1}{\epsilon}$ so that $\frac{1}{N} < \epsilon$. If $n \geq N$ then $\frac{1}{n} \leq \frac{1}{N} < \epsilon$. Thus

$$\frac{1}{n} = \left| \frac{1}{n} \right| = \left| \frac{1}{n} - 0 \right| = \left| x_n - x \right| < \epsilon$$

Since $\epsilon > 0$ was arbitrary,

$$\lim_{n \to \infty} \left(\frac{1}{n} \right) = 0 \text{ or } \left(\frac{1}{n} \to 0 \right).$$

2.2   Limits

We continue the review by considering the limits of functions. Recall that a limit is unique provided that it exists. Let $f : D \to \mathbb{R}, D \subseteq \mathbb{R}, c \in \mathbb{R}$ such that

$$[(c - \delta, c + \delta)/\{c\}] \cap D \neq \varnothing\, \forall \delta > 0.$$

Lets examine the definition of a limit of a function at some point c.

> **Definition 2.3: Limit of a function at a point.**
>
> $L \in \mathbb{R}$ is the limit of f at c if $\forall \epsilon > 0 \exists \delta > 0$ such that $|f(x) - L| < \epsilon$

whenever $|x - c| < \delta$. We write

$$\lim_{x \to c} f(x) = L \text{ or } f(x) \to L \text{ as } x \to c.$$

Note that $f(c) \neq L$ is possible. It is also not required that $c \in D$, it is only required that c is a limit point of D. Figure 2.1 depicts the limit of a function at a point. Note that

$$\lim_{x \to c} f(x) = L, \text{ however } f(x) \neq L.$$
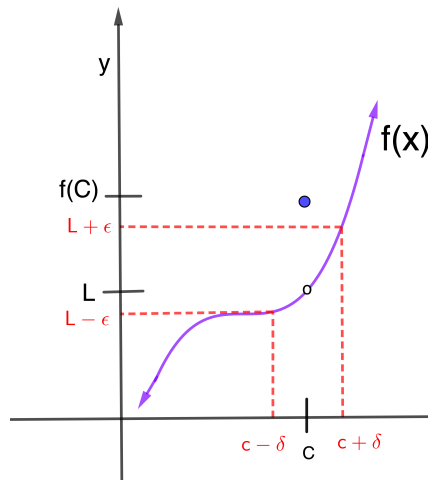
Figure 2.1: The value of ϵ is fixed first, then an interval in the domain can be found based on δ that will satisfy the predetermined value of ϵ.

Definition 2.4: Continuity of a function

Let $f : D \to \mathbb{R}$ and $c \in D$. Then f is continuous at c if

$$\lim_{x \to c} f(x) = f(c).$$

When considering the continuity of a function at a point we are looking to see that $f(x)$ can be made arbitrarily close to $f(c)$ if x is close enough to c. Continuity implies that small changes in x produce small changes in $f(x)$. The reader should think of some illustrating examples where the function f is, and is not continuous at c. How would these situations be depicted graphically?

2.3   Derivatives

Let $f : D \to \mathbb{R}$, and $a \in D$.

> **Definition 2.5: Derivative of a function at a point.**
>
> The <u>derivative of a function</u> f at a, denoted $f'(a)$ is
>
> $$f'(a) = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h} \text{ or } f'(a) = \lim x \to a \frac{f(x) - f(a)}{x - a}.$$
>
> provided the limit exists.

The following represents a list of applications of derivatives that the reader should be familiar with.

- Finding slopes of tangent lines.

- Instantaneous rate of change.

- Graphing

- Optimization

Also recall that there is a set of derivative rules and techniques including:

- Power Rule

- Product Rule

- Quotient Rule

- Chain Rule

Recall that differentiability of a function at x_0 implies continuity at x_0; however, continuity of a function at x_0 does not imply the differentiability of the function at x_0.

A key concept from calculus is the mean value theorem.

> **Theorem 2.1: The Mean Value Theorem**
>
> Suppose f is continuous on $[a, b]$, and f is also differentiable on (a, b). Then $\exists c \in (a, b)$ such that
>
> $$f'(c) = \frac{f(b) - f(a)}{b - a} \text{ or } f(b) - f(a) = f'(c)(b - a).$$

By defining $T_c(x)$ as the tangent line to $f(x)$ at c with $f'(c) = T$, and L to be the secant line between points $(a, f(a))$ and $(b, f(b))$ such that the slope between these points is given by

$$L = \frac{f(b) - f(a)}{b - a},$$

then the following figure illustrates the mean value theorem.
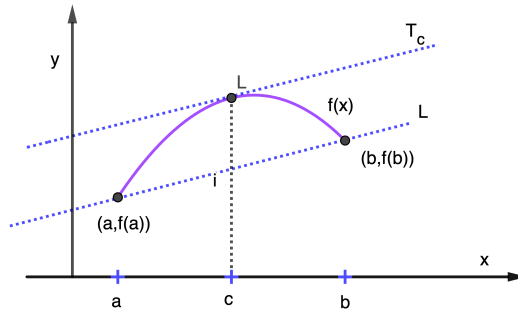
The mean value theorem assures that there exists $c \in (a, b)$ such that the slope of $T_c(x) = L$.

2.4 Antidifferentiation

Definition 2.6: Antidifferentiation

$F(x)$ is called an antiderivative of $f(x)$ on the interval I if $F'(x) = f(x) \forall x \in I$.

Note that since $F(x)$ is differentiable, it is continuous. Hence all antiderivatives are continuous. Also

$$F'(x) = f(x) \implies \frac{d}{dx}(F(x) + c) = f(x) \forall\, c \in \mathbb{R}.$$

Example 2.2: $\sin(x)$

For $f(x) = \cos(x)$. Then $F(x) = \sin(x), \sin(x) + 1$, etc.

2.5   Integration

This section can not cover all technical aspects of integration ; however, it is good to be reminded that integration comes from a limit.

Consider $f : [a, b] \to \mathbb{R}$, and partition $[a, b]$ into $n \in \mathbb{N}$ parts such that:

partition
[a,b]

x_1^* x_2^*

$a = x_0$ x_1 x_2 $b = x_n$

With
$$\Delta x_i := x_i - x_{i-1}, \forall i \in \{1, 2, \ldots, n\}, \Delta x = \max_i \Delta x_i.$$

then we explore integrable functions.

Definition 2.7: Integrable

A function f is integrable on $[a, b]$ if

$$I = \lim_{\Delta x \to 0} \sum_{i=1}^{n} f(x_i^*) \Delta x_i \text{ exists.}$$

In this case we write:

$$I = \int_{t=a}^{b} f(t) \, dt = \text{ definite integral.}$$

What types of functions will the value of I exist? Think about continuous functions, or bounded functions when discontinuities are present.

Theorem 2.2: Fundamental Theorem of Calculsu

Suppose f is continuous on $[a, b]$, Define

$$F(x) = \int_{t=a}^{x} f(t) \, dt \, \forall x \in [a, b] \qquad (2.1)$$

then
$$F'(x) = \frac{d}{dx} \int_{t=a}^{x} f(t) \, dt = f(x) \text{ on } [a, b].$$

Here $F(x)$ is an antiderivative of $f(x)$. Note that (2.1) has $F(a) = 0$ since $\Delta x_i = 0 \, \forall i$. Note that $F(x) = \int_a^x f(t) \, dt + k$, for arbitrary $k \in \mathbb{R}$, must also be an antiderivative such that $F(a) = k$. Hence,

if

$$F'(x) = f(x) \text{ on } [a, b] \text{ then } \int_{x=a}^{b} f(x)\, dx = F(b) - F(a).$$

2.6 Integration Techniques

The following are basic integration techniques that readers should be familiar with:

- **Substitution:** Used to 'reverse' chain rule .

  1. $\dfrac{d}{dx}\left(f(g(x))\right) = f'(g(x))g'(x)$

  2. $\displaystyle\int f'(g(x))g'(x)dx$ make use of $u = g(x), du = g'(x)dx$

  $$\int f'(u)du = f(u) + C = f(g(x)) + C.$$

- **Integration By Parts:** This is related to the product rule.

  1. $\dfrac{d}{dx}\left(f(x)g(x)\right) = f'(x)g(x) + g(x)g'(x)$

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

  2. $\displaystyle\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx$
     Symbolically: $u = f(x) \to du = f'(x)dx$ and $dv = g'(x)dx \to v = g(x)$

     $$\int u dv = uv - \int v du.$$

- **Trigonometric Integrals:** Know that trig identities are used to find $\int \sin^m(x)\cos^n(x)\, dx$ and $\int \tan^m(x)\sec^n(x)\, dx$.

- **Trig Substitution:** When and integral contains an expression of the form

  $$\sqrt{a^2 - x^2} \text{ then } x = \arcsin(\theta) \text{ and } 1 - \sin^2(\theta) = \cos^2(\theta).$$

  $$\sqrt{a^2 + x^2} \text{ then } x = \arctan(\theta) \text{ and } 1 + \tan^2(\theta) = \sec^2(\theta).$$

  $$\sqrt{x^2 - a^2} \text{ then } x = \operatorname{asec}(\theta)$$

Additionally know that there are techniques that are used for dealing with partial fractions, and rationalizing substitutions.

2.7   Infinite Series

Let $a_n \in \mathbb{R} \ \forall \ n \in \mathbb{N}$. The formal definition of a series is an expression of the form

$$a_1 + a_2 + \cdots + a_n + a_{n+1} + \cdots$$

with infinitely many additions. To attach meaning to a series we consider their convergence.

Definition 2.8: Series Convergence

Suppose $(a_n) \subset \mathbb{R}$ is a sequence and define

$$s_n := a_1 + a_2 + \cdots + a_n \in \mathbb{R} \ \forall \ n \in \mathbb{N}.$$

Note: $s_1 = a1$ and $s_{n+1} = s_n + a_{n+1} \ \forall \ n \geq 1$.
If (s_n) converges (say to $S \in \mathbb{R}$) then we say the series converges and set

$$\sum_{n=1}^{\infty} a_n = \lim_{n \to \infty} (s_n) = S.$$

If a series does not converge we say it diverges.

Theorem 2.3: Convergent Series

$$\sum_{n=1}^{\infty} a_n \text{ convergent } \longrightarrow \lim_{n \to \infty} a_n = 0$$

Note the converse is not true

Example 2.3: $\frac{1}{n}$ goes to zero too slow.

$$\lim_{n \to \infty} \frac{1}{n} = 0 \text{ but } \sum_{n=1}^{\infty} \frac{1}{n} = \infty \notin \mathbb{R}$$

Theorem 2.4: Absolute Convergence

$$\sum_{n=1}^{\infty} |a_n| \text{ convertent } \implies \sum_{n=1}^{\infty} a_n \text{ converges } (a_n \in \mathbb{R} \ \forall \ n \in \mathbb{N})$$

> This comes from properties of the real numbers.

Note that the converse of this theorem is not true. If

$$\sum_{n=1}^{\infty} a_n \text{ converges, and } \sum_{n=1}^{\infty} |a_n|$$

does not converge, $\sum_{n=1}^{\infty} a_n$ is conditionally convergent. If $\sum_{n=1}^{\infty} a_n$ is conditionally convergent then the terms can be rearranged and made to converge to any real number or diverge to either extreme or nothing. Consider the following example:

Example 2.4: Convergence

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = \ln(2)$$

however,

$$\sum_{n=1}^{\infty} \frac{1}{n} \text{ diverges.}$$

2.8 Convergence Tests

1. **Integral test**: This gives us information about p-series. The series

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

converges if $p > 1$ and diverges if $p \leq 1$.

Example 2.5: Geometric Series

$$\sum_{n=1}^{\infty} ar^{n-1} = \frac{a}{1-r} \text{ if } |r| < 1 \text{ and diverges if } |r| \geq 1.$$

---

2. **Comparison Test:** Given a_b and $b_n > 0 \; \forall \, n \in \mathbb{N}$:

   - $a_n \leq b_n \; \forall n \in \mathbb{N}; \sum b_n$ converges implies $\sum a_n$ converges.
   - $a_n \geq b_n \; \forall n \in \mathbb{N}; \sum b_n$ diverges implies $\sum a_n$ diverges.

3. **Limit Comparison**

4. **Alternating Series** Here we consider:

$$\sum_{n=1}^{\infty}(-1)^{n-1}b_n \text{ converges}$$

provided:

$b_n > 0$ such that $b_{n+1} \le b_n \ \forall\, n \in \mathbb{N}$ and $\lim_{n\to\infty} b_n = 0$.

5. **Ratio Test:**

$$L = \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|;$$

when

$$L < 1 \implies \text{absoloute convergence} \implies \text{convergence.}$$

when

$$L > 1 \text{ or } L = \infty \text{ divergence.}$$

6. **Root Test:**   Consider

$$L = \lim_{n\to\infty}|a_n|^{\frac{1}{n}}$$

and use the results state above.

2.9   Power Series

Define

$$f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n, a \in \mathbb{R}, c_n \in \mathbb{R} \ \forall\, n \in \mathbb{N}, \text{ and } x \text{ a variable.}$$

Theorem 2.5: Convergence of Power Series

There are 3 possibilities for the convergence of $f(x)$

1. Domain of f, $dom(f) = \{a\}$ converges only at $x = a$.

2. $dom(f) = \mathbb{R}$ converges $\ \forall\, x \in \mathbb{R}$.

3. The series converges for $x \in (a-R, a+R)$ and diverges if $|x - a| > R$. The domain may or may not converge conditionally at $x = a \pm R$.

Power series converge uniformly in their interval of convergence and hence are

1. Continuous

2. Differentiable (term-by-term)

3. Integrable (term-by-term).

If f is defined in some interval by a power series, f is called analytic.

2.10   Exercises

Work the following exercises:

1. Show that $\lim_{n \to \infty} \left(\frac{n}{n+1} \right) = 1$.

2. For the function
$$f(x) = \begin{cases} x^2 - 3 & x \leq 2 \\ x - 1 & x > 2, \end{cases}$$
show that f is continuous at $x = 2$. Hint: consider 1-sided limits and recall that $\lim_{x \to c} p(x) = p(c)$, for p a polynomial, and $c \in \mathbb{R}$.

3. Explain why f is not continuous at any point in $\mathbb{R}$ for the function
$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R}/\mathbb{Q}. \end{cases}$$

4. For $f(x) = x^2 - 8x + 9$ find $f'(x)$ using the limit definition.

5. Given $k(x) = \sqrt{4 - x^2} e^{7x}$. Find $k'(x)$ and equation of a tangent line at $x = 0$. Write a sentence interpreting the meaning of $k'(0)$.

6. Given $f(x) = |x|$. Show f is continuous at $x = 0$, but not differentiable at $x = 0$.

7. Use the mean value theorem to explain why $x^5 + 10x + 3 = 0$ has exactly one solution.

8. Given,
$$f(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0. \end{cases}$$
explain why $f(x)$ has no antiderivative on any interval containing 0.

9. Evaluate:

(a) $\displaystyle\int_{x=e}^{e^4} \frac{1}{x\sqrt{\ln(x)}}\, dx$

(b) $\displaystyle\int e^x \sin(x)\, dx$

(c) Given the power series

$$f(x) = \sum_{n=0}^{\infty} \frac{(x-3)^n}{n}.$$

Find the domain of f, that is values of x where the series converges. Consider the end points of the interval you find.

Chapter 3

Ordinary Differential Equations

Mathematics is a universal language and is progressed because people attempt to provide answers to questions human beings have about the world. This chapter gives some historical context and a brief review of some concepts from ordinary differential equations. The study of differential equations is necessary to advance our study and assist in answering questions.

3.1   History and Background

Calculus and ordinary differential equations are based in the work of prominent mathematicians, physicists, philosophers and scientists including:

- Isaac Newton: Physicist, Philosopher

- Gottfried Wilhelm Leibniz: Philosopher, Scientist

- Leonhard Euler: Physicist, Astronomer

According to some mathematics historians, the study of differential equations began in 1675 when Gottfried Leibniz wrote

$$\int \mathbf{x} d\mathbf{x} = \frac{1}{2}\mathbf{x}^2.$$

Isaac Newton had the essence of the method of fluxions (equivalent to Leibniz's differential and integral calculus) by 1666. His method of

integration by infinite series became known to other mathematicians privately in 1668. He also delivered them in Cambridge between 1673 to 1683. They were finally published in 1707.

The devil is in the detail, but don't forget the big picture when you are working on the detail.

3.2   Basic Concepts and Definitions

What is derivative?

- A differential equation involving only ordinary derivatives with respect to a single independent variable is called an **ordinary differential equation, ODE.** A differential equation involving partial derivatives with respect to more than one independent variable is a **partial differential equation, PDE.**

- The **order** of a differential equation is the order of the highest-order derivatives present in the equation.

Definition 3.1: differential Equation

A differential equation is **linear** if it has the format

$$a_n(x)\frac{d^n y}{dx^n} + a_{n-1}(x)\frac{d^{n-1}y}{dx^{n-1}} + \cdots + a_1(x)\frac{dy}{dx} + a_0(x)y = F(x),$$

where $a_n(x), a_{n-1}(x), \ldots, a_0(x)$, and $F(x)$ depend only on the independent variable x.

Example 3.1: ODE or PDE

Determine each of the following equations is ODE or PDE, linear or nonlinear, and the order.

1. $\frac{d^2 y}{dx^2} + y^3 = x^2$
   Note differentiation is with respect to only x.

2. $\frac{d^3 y}{dx^3} + \cos(x)\frac{d^2 y}{dx^2} - \frac{dy}{dx} + y = \frac{x^2+1}{x-2}$
   Again note differentiation is with respect to only x.

3. $\frac{\partial^4 u}{\partial y^4} - \frac{\partial^3 u}{\partial x^3} + u^2 = x^3 - y^3$.
   Here their is differentiation with respect to both x and y.

Definition 3.2: nth-order ODE

A general form for an nth-order equation with x independent, y dependent can be expressed as

$$F\left(x, y, \frac{dy}{dx}, \ldots, \frac{d^n y}{dx^n}\right) = 0, \qquad (3.1)$$

where F is a function that depends on x, y and the derivatives of y up to order n. It is assumed that the equation holds for all x in an open interval $I(a < x < b$, where a or b could be infinite). In many cases (3.1) can be written as

$$\frac{d^n y}{dx^n} = f\left(x, y, \frac{dy}{dx}, \ldots, \frac{d^{n-1} y}{dx^{n-1}}\right), \qquad (3.2)$$

which is often preferable to (3.1).

Definition 3.3: Initial Value Problem

A **Initial Valued Problem** is a differential equation along with appropriate number of initial conditions. That is, the equation (3.1) on an interval I satisfies at x_0 the n initial conditions

$$\begin{aligned}
y(x_0) &= y_0, \\
\frac{dy}{dx}(x_0) &= y_1, \\
&\vdots \\
\frac{d^{n-1} y}{dx^{n-1}}(x_0) &= y_{n-1},
\end{aligned}$$

where $x_0 \in I$ and $y_0, y_1, \ldots, y_{n-1}$ are given constants.

Any ordinary differential equation can be written as a first-order ordinary differential equation.

Example 3.2: ODE Example

Given $y'' - y' - 2y = 0$. Rewrite it as a 1st order ODE. Show that $\phi(x) = c_1 e^{-x} + c_2 e^{2x}$ is a solution to the given ODE for any choice of the constants c_1 and c_2. Finally determine c_1 and c_2 so that the initial conditions $y(0) = 2$ and $y'(0) = -1$.

Without (initial/boundary) conditions, an ODE is likely to have **inifite** solutions. **WHY??**

Definition 3.4: Direction Field

A plot of short line segments drawn at various points in the xy-plane showing the slop of the solution curve is called a **direction field** for the differential equation.

Example 3.3: Direction Field

Given the direction field of $y' = x^2 - y$. Sketch the solution for each of the following initial conditions.

1. $y(-.5) = 1$ 2. $y(-1) = 0$ 3. $y(0) = 1$
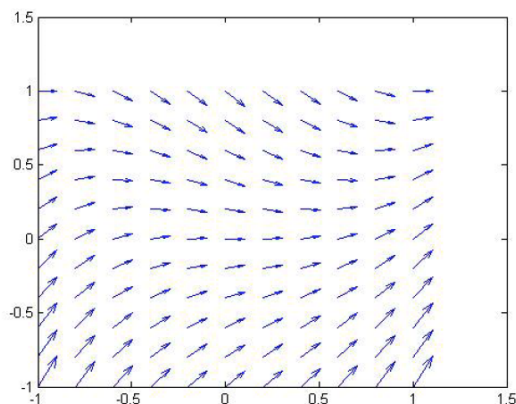
Matlab is sometimes useful to assist in plotting the direction field.

Example 3.4: Matlab Plotting

```
[x,y] = meshgrid(-1:.2:1,-1:.2:1);
dx=ones(size(x));
dy=x.^2-y;
quiver(x,y,dx,dy)
```

Use the above matlab code to assist yourself in plotting the direction field for each of the following differential equations.

1. $y' = \frac{1}{x - .5x^2}$ over $[0, 12] \times [0, 12]$ with the step size of .5.

2. $y' = (y^2-y-2)(1-y)^2$ over $[0,5] \times [-2,3]$ with the appropriate step size.

What does the `.^` command do?

3.3 First Order Differential Equations

Lets consider solving some differential equations.

> **Definition 3.5: Separable**
>
> A **separable** differential equation is any DE that can be written in the form of
> $$\frac{dy}{dx} = g(x)p(y). \tag{3.3}$$

To solve a separable differential equation, we can rewrite (3.3) as

$$h(y)dy = g(x)dx, \text{ where } h(y) = \frac{1}{p(y)}$$

. Then integrate both sides

$$\int h(y)dy = \int g(x)dx \Rightarrow H(y) = G(x) + C,$$

where we merged the two constants of integration into a single symbol C. The last equation gives an **implicit** solution to (3.3).

> **Example 3.5: Separable Equation**
>
> Given $\frac{dy}{dx} = \frac{x-5}{y^2}$ with $y(0) = 12$.
>
> 1. Plot a direction field for the given DE and sketch the solution of the IVP.
>
> 2. Find the solution to the IVP algebraically.
>
> 3. Plot the solution from part 2 on the direction field.

3.4 Linear Equations

> **Definition 3.6: Linear 1st Order Equation**
>
> A **linear 1st order equation** is an equation that can be expressed

in the form

$$a_1(x)\frac{dy}{dx} + a_0(x)y = b(x), \qquad (3.4)$$

where $a_1(x)$, $a_0(x)$, and $b(x)$ depend only on the independent variable x, not on y. The **standard form** of a linear 1st order DE is

$$\frac{dy}{dx} + P(x)y = Q(x). \qquad (3.5)$$

Some questions for thought:

1. Is it possible that (3.4) is separable?

2. What is the advantage of the standard form (3.5)?

Solving linear equations like (3.4), the idea is if $a_0(x) = a_1'(x)$, then (3.4) becomes

$$\frac{d}{dx}[a_1(x)y] = b(x) \Rightarrow a_1(x)y = \int b(x)dx \Rightarrow y = \frac{1}{a_1(x)}\int b(x)dx.$$

However, it is very unlikely that $a_0(x) = a_1'(x)$. If considering (3.5) and multiplying $\mu(x)$ to (3.5), i.e.,

$$\mu(x)\frac{dy}{dx} + \mu(x)P(x)y = \mu(x)Q(x).$$

If we choose $\mu(x)$ such that $\mu'(x) = \mu(x)P(x)$. Then the left hand side become $\frac{d}{dx}[\mu(x)y]$.

1. Define $\mu(x) = e^{\int P(x)dx}$.

2. Integrate $\frac{d}{dx}[\mu(x)y] = \mu(x)Q(x)$.

3. Solve for y by dividing by $\mu(x)$.

3.5   Exact Equations

Definition 3.7: Exact Equations

The differential from $M(x,y)dx + N(x,y)dy$ is said to be **exact** in a rectangle R if there is a function $F(x,y)$ such that

$$\frac{\partial F}{\partial x}(x,y) = M(x,y) \text{ and } \frac{\partial F}{\partial y}(x,y) = N(x,y)$$

for all (x,y) in R. That is, the total differential of $F(x,y)$ satisfies $dF(x,y) = M(x,y)dx + N(x,y)dy$. If $M(x,y)dx + N(x,y)dy$ is an

exact differential form, then the equation

$$M(x, y)dx + N(x, y)dy = 0 \qquad (3.6)$$

is called an **exact equation**. Its solutions are given by the level curves $F(x, y) = c$, for an arbitrary constant c.

To solve an exact equation:

1. Verify that it is an exact equation: $\frac{\partial F}{\partial x} = M$ and $\frac{\partial F}{\partial y} = N$.

2. Define $F(x, y) = \int M(x, y)dx + g(y)$

3. Take partial derivative w.r.t y of both sides of Step 2 and substitute N for $\frac{\partial F}{\partial y}$. Then solve for $g'(y)$.

4. Integrate $g'(y)$ to obtain $g(y)$ and substitute $g(y)$ into Step 2.

Example 3.6: Exact Equation

Given $\frac{dy}{dx} = -\frac{2xy^2 + 1}{2x^2 y}$.

1. Solve the above differential equation.

2. Plot level curves of the implicit solution.

3.6 Special Integrating Factors

If (3.6) is not exact, is it possible to find an equivalent equation which is exact?

Consider the techniques from the section on linear equations, is it possible to choose $\mu(x, y)$ such that

$$\mu(x, y)M(x, y)dx + \mu(x, y)N(x, y)dx = 0 \text{ is exact?}$$

If $\mu(x, y)$ exists, we call $\mu(x, y)$ an **integrating factor** of (3.6).

3.7 Substitutions and Transformations

Another common technique used when solving differential equations is to consider substitutions and transformations. Here the procedure is:

1. Identify the type of equation and determine the appropriate substitution or transformation.

2. Rewrite the original equation in terms of new variables.

3. Solve the transformed equation.

4. Express the solution in terms of the original variables.

Example 3.7: Transformation

Solve $\frac{dy}{dx} = y - x - 1 + (x - y + 2)^{-1}$.

Transformations are very common techniques in dealing with "unsolvable" problems. The main idea is moving the problem to a different "space" that makes the problem easier to solve and then move the solution back to the original space.

3.8   Interval of Validity

Does the solution exist for an initial value problem? And if so, is the solution unique for a given IVP?

Theorem 3.1: Unique Solution on Interval for IVP

Given the initial value problem (IVP)

$$\frac{dy}{dx} = f(x, y), y(x_0) = y_0,$$

assume that f and $\frac{\partial f}{\partial y}$ are continuous function in rectangle $R = \{(x, y) : a < x < b, c < y < d\}$ that contains the point (x_0, y_0). Then the IVP has a unique solution $\phi(x)$ in some interval $x_0 - \delta < x < x_0 + \delta$, where $\delta > 0$.

Example 3.8: Interval of Validity

For the initial value problem

$$\frac{dy}{dx} = 3y^{2/3}, y(2) = 0.$$

1. Does Theorem 3.1 imply the existence of a unique solution?

2. Find two solutions for the given IVP.

3. Consider the initial condition, $y(0) = 10^{-7}$, does it have a

> unique solution in the neighborhood of $x = 0$?

3.9   Euler's Method

Euler's method (or the tangent-line method): constructs approximate solutions to an IVP for a 1st order ODE of the form:

$$y' = f(x, y), y(x_0) = y_0. \tag{3.7}$$

If (3.7) has a unique solution $\phi(x)$ in the neighborhood of x_0, i.e. $\phi'(x) = f(x, \phi(x)), \phi(x_0) = y_0$. Let $x_n = x_0 + nh$, $n = 0, 1, 2, \ldots$, where $h > 0$ is the step size. How to approximate $\phi(x_1)$, $\phi(x_2), \ldots$, if the explicit solution cannot be found?

Theorem 3.2: Taylor's Theorem

If $\phi : [a, b] \to \mathbb{R}$ is $n+1$-times continuously differentiable, then there exists $c \in (a, b)$ such that

$$\phi(b) = \phi(a) + \phi'(a)(b-a) + \cdots + \frac{\phi^{(n)}(a)}{n!}(b-a)^n + \frac{\phi^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1}.$$

Assume that $\phi(x)$ is twice continuously differentiable. According to Theorem 3.2,

$$\phi(x) \approx \phi(x_0) + \phi'(x_0)(x - x_0), \text{ where } x \in \text{neiborhood of } x_0.$$

If h is small enough, we can derive the followings:

$$\begin{aligned}
\phi(x_1) &\approx \phi(x_0) + \phi'(x_0)(x_1 - x_0) = y_0 + hf(x_0, y_0) \\
\phi(x_2) &\approx \phi(x_1) + \phi'(x_1)(x_2 - x_1) = \phi(x_1) + hf(x_1, \phi(x_1)) \\
\phi(x_3) &\approx \phi(x_2) + \phi'(x_2)(x_3 - x_2) = \phi(x_2) + hf(x_2, \phi(x_2))
\end{aligned}$$

Denote $y_0 + hf(x_0, y_0)$ as y_1, which is an approximation of $\phi(x_1)$. Thus,

$$\phi(x_2) \approx \phi(x_1) + hf(x_1, \phi(x_1)) \approx y_1 + hf(x_1, y_1).$$

Similarly, we define $y_{k+1} = y_k + hf(x_k, y_{k+1})$.

Then we can state **Euler Method:**

$$\begin{aligned}
x_{n+1} &= x_n + h \\
y_{n+1} &= y_n + hf(x_n, y_n), n = 0, 1, 2, \ldots.
\end{aligned}$$

Example 3.9: Euler's Method

Consider the IVP $y' + 2y = 2 - e^{-4x}, y(0) = 1$. We know the solution is $y(x) = 1 + .5e^{-4x} - .5e^{-2x}$. Compare the results from Euler Method with the solution using different h for the interval of $[0, 5]$.
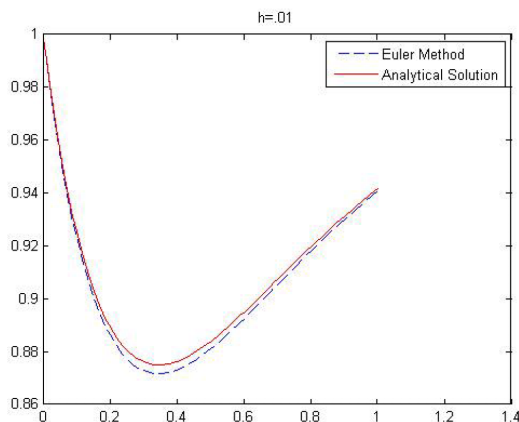
A matlab code that can be used to find the solution to the given IVP is given by:

```matlab
f=inline('2-exp(-4*x)-2*y','x','y');
x0=0;
y0=1;
x=[x0];
y=[y0];
h=.01;
n=100;
for i=1:n
    x1=x0+h;
    y1=y0+h*f(x0,y0);
    x=[x;x1];
    y=[y;y1];
    x0=x1;
    y0=y1;
end

sol=inline('1+.5*exp(-4*x)-.5*exp(-2*x)','x');
soly=sol(x);
plot(x,y,'b--',x,soly,'r'), legend('Euler Method','Analytical Solution')
```

3.10   Second order DE with constant coefficients

> **Definition 3.8: Second Order Constant Coefficient DE**
>
> A second order constant coefficient differential equation is in the form of
> $$ay'' + by' + cy = f(t), a \neq 0 \qquad (3.8)$$
> with the special case where the function $f(t)$ is 0:
> $$ay'' + by' + cy = 0, \qquad (3.9)$$
> is the homogeneous form of (3.8).

Can the reader convert (3.8) and (3.9) to 1st order ODE?

From (3.9), we observe that y'' is a linear combination of y' and y. What elementary functions have this property?

$$y(t) = e^{rt}.$$

If we substitute $y = e^{rt}$ into (3.9), we obtain

$$ar^2 e^{rt} + bre^{rt} + ce^{rt} = 0 \Rightarrow e^{rt}(ar^2 + br + c) = 0.$$

Since $e^{rt} \neq 0, \forall t$, the above equation is equivalent to

$$ar^2 + br + c = 0. \qquad (3.10)$$

Here (3.10) is known as the auxiliary equation or characteristic equation . Think about what happens to the equivalent system of 1st order ODEs?

> **Theorem 3.3: Principle of Superposition**
>
> If $y_1(t)$ and $y_2(t)$ are two solutions to a linear, homogeneous differential equation, then so is $y(t) = c_1 y_1(t) + c_2 y_2(t)$.

> **Example 3.10**
>
> olve the initial value problem
> $$y'' + 3y' - 1y = 0; y(0) = 0, y'(0) = -1;$$

3.11 Method of Undetermined Coefficients

To find a particular solution to the differential equation

$$ay'' + by' + cy = P_m(t)e^{\alpha t}\cos \beta t + Q_n(t)e^{\alpha t}\sin \beta t,$$

where $P_m(t)$ and $Q_n(t)$ are polynomials of degree m and n, respectively. Use the form

$$
\begin{aligned}
y_p(t) \quad = \quad & t^s(a_k t^k + \cdots + a_1 t + a_0)e^{\alpha t}\cos\beta t \\
& +t^s(b_k t^k + \cdots + b_1 t + b_0)e^{\alpha t}\sin\beta t,
\end{aligned}
$$

where k is the larger of m and n. If $\alpha + i\beta$ is not a root of the associated characteristics equation, take $s = 0$; if $\alpha + i\beta$ is a root of the associated characteristics equation, take $s = 1$.

Example 3.11: Undetermined Coefficients

Use the method of undetermined coefficients to solve:

1. Solve $y'' - y = 2 - t^2, y(0) = 1, y'(0) = 0$.

2. Solve $36y'' - 12y' + 37y = 0$.

3.12   Eigenvalues and Eigenvectors

To solve a system of 1st order ODEs, eigenvalues and eigenvectors are used frequently.

Definition 3.9: Eigenvalue and Eigenvector

Given A is a $n \times n$ matrix. λ is an eigenvalue of A if there exists a non zero vector $\mathbf{x}$ such that

$$
A\mathbf{x} = \lambda\mathbf{x}.
$$

$\mathbf{x}$ is called a eigenvector associated with the eigenvalue λ.

Example 3.12: Eigenvalue and Eigenvector

Find eigenvalues and associated eigenvectors of $A = \begin{bmatrix} -1 & 8 \\ 2 & -1 \end{bmatrix}$.

3.13   Exercises

1. Consider $y' = (y^2 - y - 2)(1 - y)^2$ and use the information from Example (3.4) to describe the behavior of solutions with different initial conditions. Note that matlab quiver may be helpful when plotting the direction field.

2. According to Newton's law of cooling, if an object at temperature T is immersed in a medium having the constant temperature M, the rate of change of T is proportional to the difference of temperature $M - T$.

   (a) Write an equation that describes the Newton's law of cooling?

   (b) Solve the equation.

   (c) A thermometer reading 100^oF is placed in a medium having a constant temperature of 70^oF. After 6 minutes, the thermometer reads 80^oF. What is the reading after 20 minutes?

3. A rock contains two radioactive isotopes, RA_1 and RA_2, the belong to the same radioactive series; that is RA_1 decays into RA_2, which then decays into stable atoms. Assume that the rate at which RA_1 decays into RA_2 is $50e^{-10t}kg/sec$. Because the rate of decay of RA_2 is proportional to the mass $y(t)$ of RA_2 present.

   (a) Write an differential equation describing the rate of change in RA_2 using the notations: the mass $y(t)$ of RA_2 and the decay constant of $k > 0$.

   (b) If $k = 2/sec$, and the initial mass of $40kg$, find the mass $y(t)$ of RA_2 for $t \geq 0$.

   (c) Approximate how long it takes for RA_2 to be less than 0.1kg.

4. Consider the IVP $y' - y = -0.5e^{0.5x}\sin(5x) + 5e^{0.5x}\cos(5x), y(0) = 0$. Its solution is $y(x) = e^{0.5x}\sin(5x)$. Use Euler's method to find the approximated solutions with $h = 0.1, 0.05, 0.001, 0.0001$ and compare with the analytic solution.

5. Convert the given IVP to a system of 1st order ODEs and solve the system.
$$y'' + 3y' - 1y = 0; y(0) = 0, y'(0) = -1;$$

6. Use the method of undetermined coefficients to solve:

   (a) Solve $36y'' - 12y' + 37y = 3e^{t/6}\cos t$.

   (b) Solve $36y'' - 12y' + 37y = 5e^{-2t}\sin 2t$.

Chapter 4

Linear Algebra

A brief review of concepts from linear algebra.

4.1 Row Reduction and Solving Systems of Equations

The three elementary row operations are:

$Ri \leftrightarrow Rj$ Swap Row i with Row j.

mRi Replace Row i with m(Row i), where $m \neq 0$.

$mRi + Rj$ Replace Row j by the sum of itself and m(Row i).

We use elementary row operations to find an equivalent matrix in reduced row echelon form, RREF, (or even just row echelon form , REF) to solve systems of equations.

Example 4.1: Row Reduction

Solve the following system using elementary row operations:

$$
\begin{array}{rcrcrcrcr}
x_1 & + & 3x_2 & - & 2x_3 & + & 5x_4 & = & 4 \\
2x_1 & + & 8x_2 & - & x_3 & + & 9x_4 & = & 9 \\
3x_1 & + & 5x_2 & - & 12x_3 & + & 17x_4 & = & 7
\end{array}
$$

Solution: Convert to an augmented matrix , and then use elementary row operations to find an equivalent matrix in either REF or

---

RREF.

$$
\begin{bmatrix}
1 & 3 & -2 & 5 & 4 \\
2 & 8 & -1 & 9 & 9 \\
3 & 5 & -12 & 17 & 7
\end{bmatrix}
\xrightarrow[\substack{-3R1+R3}]{\substack{-2R1+R2}}
\begin{bmatrix}
1 & 3 & -2 & 5 & 4 \\
0 & 2 & 3 & -1 & 1 \\
0 & -4 & -6 & 2 & -5
\end{bmatrix}
$$

$$
\xrightarrow{0.5R2}
\begin{bmatrix}
1 & 3 & -2 & 5 & 4 \\
0 & 1 & 1.5 & -0.5 & 0.5 \\
0 & -4 & -6 & 2 & -5
\end{bmatrix}
$$

$$
\xrightarrow[\substack{4R2+R3}]{\substack{-3R2+R1}}
\begin{bmatrix}
1 & 0 & -6.5 & 6.5 & 2.5 \\
0 & 1 & 1.5 & -0.5 & 0.5 \\
0 & 0 & 0 & 0 & -3
\end{bmatrix}
$$

This matrix is in REF, but we can still read off the solution. Since the last row of the matrix is equivalent to the equation $0 = -3$, the system has no solution (i.e. is inconsistent).

Example 4.2: Elementary Row Operations

Solve the following system using elementary row operations:

$$
\begin{array}{rrrrrrr}
x & + & y & - & z & = & 0 \\
2x & - & 3y & + & z & = & 0 \\
x & - & 4y & + & 2z & = & 0
\end{array}
$$

Solution: Convert to an augmented matrix, and then use elementary row operations to find an equivalent matrix in either REF or RREF.

$$
\begin{bmatrix}
1 & 1 & -1 & 0 \\
2 & -3 & 1 & 0 \\
1 & -4 & 2 & 0
\end{bmatrix}
\xrightarrow[\substack{-R1+R3}]{\substack{-2R1+R2}}
\begin{bmatrix}
1 & 1 & -1 & 0 \\
0 & -5 & 3 & 0 \\
0 & -5 & 3 & 0
\end{bmatrix}
$$

$$
\xrightarrow{-R2+R3}
\begin{bmatrix}
1 & 1 & -1 & 0 \\
0 & -5 & 3 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
\xrightarrow{-0.2R2}
\begin{bmatrix}
1 & 1 & -1 & 0 \\
0 & 1 & -0.6 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
$$

This final matrix, which is in RREF, translates to the system:

$$\begin{aligned} x &= 0.4z \\ y &= 0.6z \\ z &= z \end{aligned}$$

So, z is a free variable, and the solution set is

$$\{(0.4z, 0.6z, z) \mid z \in \mathbb{R}\}.$$

4.2   Linear independence

A set of vectors $\{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_p}\}$ in $\mathbb{R}^n$ is said to be *linearly independent* if the vector equation

$$x_1\vec{v_1} + x_2\vec{v_2} + \cdots + x_p\vec{v_p} = \vec{0}$$

has only the trivial solution. The set is called *linearly dependent* otherwise.

Example 4.3: Linear Independent

Determine if the following set is linear independent or linearly dependent in $\mathbb{R}^3$:

$$\left\{ \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}.$$

Solution: We need to see if the vector equation

$$x_1 \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

has a nontrivial solution or if it only has the trivial solution. Forming the associated augmented matrix and row reducing, we obtain:

$$\begin{bmatrix} 2 & 1 & -3 & 0 & 0 \\ 4 & 2 & 0 & -1 & 0 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 0 & 7/6 & 0 \\ 0 & 1 & 0 & -17/6 & 0 \\ 0 & 0 & 1 & -1/6 & 0 \end{bmatrix}.$$

Since there are nontrivial solutions (notice the free variable), the set of vectors is linearly dependent.

This example illustrates that in $\mathbb{R}^n$ if a set of vectors has more than n vectors the set will necessarily be linearly dependent.

Example 4.4: Linearly Independent

Determine if the following set is linear independent or linearly dependent in $\mathbb{R}^3$:

$$\left\{ \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 7 \\ 1 \end{bmatrix} \right\}.$$

Solution: Jumping right to the augmented matrix and row reducing, we get:

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ -3 & 0 & 7 & 0 \\ 4 & 1 & 1 & 0 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Since there is only the trivial solution, the set of vectors is linearly independent.

4.3   Special classes of matrices

The *main diagonal* of a matrix refers to the entries in positions (i, i). From this we can define a *diagonal matrix* to be a matrix that has zeroes in all positions (i, j) where $i \neq j$. We usually use the classification of *diagonal matrix* for square matrices.

Example 4.5: Diagonal Matrices

Here are a couple of examples of diagonal matrices:

$$\begin{bmatrix} -1 & 0 \\ 0 & 8 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 7 \end{bmatrix}.$$

Definition 4.1: Identity Matrices

The *identity matrices* are comprised of the $n \times n$ matrices whose (i, i)-entries are 1 and whose other entries are 0.

Note that identity matrices are just a special class of diagonal matrices

> **Example 4.6: Identity Matrices**
>
> Here are some of the identity matrices:
>
> $$I_1 = [1] \quad , \quad I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad , \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad , \quad \ldots$$

An *upper triangular matrix* is a matrix that has entries of zero below the main diagonal.

> **Example 4.7: Upper Triangular**
>
> Here are two examples of upper triangular matrices:
>
> $$\begin{bmatrix} 1 & -2 & 0 \\ 0 & 2 & 9 \\ 0 & 0 & -3 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 4 & -1 \\ 0 & -2 & 7 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 8 \end{bmatrix}.$$

Similarly, a *lower triangular matrix* is a matrix that has entries of zero above the main diagonal.

> **Example 4.8: Lower Triangular**
>
> Here are two examples of lower triangular matrices:
>
> $$\begin{bmatrix} 2 & 0 & 0 \\ -1 & 0 & 0 \\ 8 & 3 & 4 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -5 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 3 & -4 & 0 \\ 2 & 0 & -1 & 4 \end{bmatrix}.$$

4.4 Matrix operations

Consider the following matrices:

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 0 \\ 4 & -1 \end{bmatrix}, B = \begin{bmatrix} 5 & 0 \\ 0 & -2 \\ 3 & 1 \end{bmatrix}, C = \begin{bmatrix} 4 & -2 & 0 \\ 1 & 0 & -1 \end{bmatrix}, D = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$$

A scalar multiple of a matrix is just multiplying every entry of the matrix by that scalar.

Example 4.9: Scalar Multiplication

Consider the matrix C. Then

$$4C = \begin{bmatrix} 4(4) & 4(-2) & 4(0) \\ 4(1) & 4(0) & 4(-1) \end{bmatrix} = \begin{bmatrix} 16 & -8 & 0 \\ 4 & 0 & -4 \end{bmatrix}$$

To add or subtract two matrices, the matrices must be the same size; and if so, the addition (or subtraction) is done component-wise.

Example 4.10: Size Matters

$A + C$ does not exist, because the matrices have different sizes (A is a 3×2 matrix, whereas C is a 2×3 matrix). However,

$$A + B = \begin{bmatrix} 2+5 & 3+0 \\ 1+0 & 0+(-2) \\ 4+3 & -1+1 \end{bmatrix} = \begin{bmatrix} 7 & 3 \\ 1 & -2 \\ 7 & 0 \end{bmatrix}$$

and

$$A - B = \begin{bmatrix} 2-5 & 3-0 \\ 1-0 & 0-(-2) \\ 4-3 & -1-1 \end{bmatrix} = \begin{bmatrix} -3 & 3 \\ 1 & 2 \\ 1 & -2 \end{bmatrix}$$

The *transpose* of a matrix is formed by switching the rows and the columns of the matrix. If the matrix is M, then the transpose is denoted by M^{T}.

Example 4.11: Transpose

Consider the matrix A. Then

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

The *product* of two matrices, M and N, can only be done if the number of columns of matrix M is the same as the number of rows of matrix N. Furthermore, the (i, j)-th entry of MN is the dot product of row i of matrix M with column j of matrix N. If M has size $m \times n$ and N has size $n \times p$, then MN will have size $m \times p$.

Example 4.12: Multiplication

Notice that AB does not exist, because A has 2 columns, whereas

B has 3 rows. However, we can compute AC.

$$AC = \begin{bmatrix} 2(4) + 3(1) & 2(-2) + 3(0) & 2(0) + 3(-1) \\ 1(4) + 0(1) & 1(-2) + 0(0) & 1(0) + 0(-1) \\ 4(4) + (-1)(1) & 4(-2) + (-1)(0) & 4(0) + (-1)(-1) \end{bmatrix}$$

$$= \begin{bmatrix} 11 & -4 & -3 \\ 4 & -2 & 0 \\ 15 & -8 & 1 \end{bmatrix}$$

It should also be noted that matrix multiplication is not commutative. In the above example, $AC \neq CA$. In fact, CA doesn't even exist!!

A matrix M is said to be *invertible*, if there exists a matrix N such that $MN = I$ and $NM = I$. If such a matrix exists, then N is called the *inverse* of M and is denoted by M^{-1}. From this definition, you can see that the only matrices that could possibly be invertible are square matrices. However, not all square matrices are invertible. One way to find the inverse of a $n \times n$ square matrix M is to form the matrix $[M \mid I_n]$ and find the row equivalent matrix in RREF. If you can get the row equivalent matrix in the form $[I_n \mid B]$, then $B = M^{-1}$. If you can't get the left-side to be I_n, then M^{-1} does not exist.

Example 4.13: Matrix Inverse

Determine A^{-1}, if it exists, for

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

Solution: Form the matrix $[A \mid I_2]$ and row reduce:

$$\begin{bmatrix} 2 & 4 & 1 & 0 \\ -1 & -2 & 0 & 1 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

Thus, A^{-1} does not exist.

Example 4.14: Matrix Inverse

Determine B^{-1}, if it exists, for

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

<u>Solution</u>: Form the matrix $[A \mid I_2]$ and row reduce:

$$
\left[\begin{array}{rrrrrr}
1 & 2 & -1 & 1 & 0 & 0 \\
-2 & 0 & 1 & 0 & 1 & 0 \\
1 & -1 & 0 & 0 & 0 & 1
\end{array}\right]
\xrightarrow{RREF}
\left[\begin{array}{rrrrrr}
1 & 0 & 0 & 1 & 1 & 2 \\
0 & 1 & 0 & 1 & 1 & 1 \\
0 & 0 & 1 & 2 & 3 & 4
\end{array}\right].
$$

Then,

$$
B^1 = \left[\begin{array}{rrr}
1 & 1 & 2 \\
1 & 1 & 1 \\
2 & 3 & 4
\end{array}\right].
$$

You can use inverses of matrices to solve systems of equations, where the coefficient matrix is a square matrix.

Example 4.15: Inverse for Solutions

Solve the following system:

$$
\begin{array}{rrrrr}
x & + & 2y & - & z & = & 4 \\
-2x & & & + & z & = & 1 \\
x & - & y & & & = & -5
\end{array}.
$$

<u>Solution</u>: As a matrix equation, this system becomes:

$$
\left[\begin{array}{rrr}
1 & 2 & -1 \\
-2 & 0 & 1 \\
1 & -1 & 0
\end{array}\right]
\cdot
\left[\begin{array}{r}
x \\
y \\
z
\end{array}\right]
=
\left[\begin{array}{r}
4 \\
1 \\
-5
\end{array}\right].
$$

Notice that the coefficient matrix is B from the previous example. So, the matrix equation is:

$$
B \cdot \left[\begin{array}{r}
x \\
y \\
z
\end{array}\right]
=
\left[\begin{array}{r}
4 \\
1 \\
-5
\end{array}\right].
$$

Then,

$$
B^{-1}B \cdot \left[\begin{array}{r}
x \\
y \\
z
\end{array}\right]
= B^{-1} \cdot
\left[\begin{array}{r}
4 \\
1 \\
-5
\end{array}\right].
$$

So,

$$
\left[\begin{array}{r}
x \\
y \\
z
\end{array}\right]
=
\left[\begin{array}{rrr}
1 & 1 & 2 \\
1 & 1 & 1 \\
2 & 3 & 4
\end{array}\right]
\cdot
\left[\begin{array}{r}
4 \\
1 \\
-5
\end{array}\right]
=
\left[\begin{array}{r}
-5 \\
0 \\
-9
\end{array}\right].
$$

4.5   Determinants

The *determinant* of a square matrix M, denoted $\det(M)$ or $|M|$, is a scalar value associated to the matrix that can be used to provide information about the matrix, including analyzing solution sets to systems of equations, finding the inverse of a matrix, finding eigenvalues, solving differential equations, etc. There are several methods for determining the determinant of a matrix, but we will focus on the cofactor expansion method. We start by providing the formula for the determinant of a 2×2 matrix:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Example 4.16: Determinants

Find the determinant of

$$\begin{vmatrix} 2 & 7 \\ -5 & 3 \end{vmatrix}.$$

<u>Solution</u>: Using the formula provided above:

$$\begin{vmatrix} 2 & 7 \\ -5 & 3 \end{vmatrix} = (2)(3) - (-5)(7) = 6 + 35 = 41.$$

We will use the determinants of these 2×2 matrices as building blocks to find the determinants of larger matrices. Associated to any square matrix is a sign matrix, which is a matrix whose entries are alternating "+" or "-" signs, starting with a "+" sign in the upper left entry:

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Then, you can use this to find determinants of larger square matrices.

Definition 4.2: Determinant

The *determinant* of an $n \times n$ matrix A is:

$$\det(A) := a_{11} \det(A_{11}) - a_{12} \det(A_{12}) + \cdots + (-1)^{1+n} a_{1n} \det(A_{1n}).$$

> where A_{ij} denotes the $(n-1) \times (n-1)$ matrix formed from A by deleting row i and column j.

If you notice, this is just an expansion along the first row of the matrix A, with the alternating signs corresponding to the entries of the first row of the sign matrix. This means that you generalize this idea to find a determinant by expanding upon any row or any column of the matrix.

Example 4.17: Determinants

Compute:
$$\det \left(\begin{bmatrix} 3 & -1 & 2 \\ 1 & 4 & 0 \\ -2 & 5 & 1 \end{bmatrix} \right)$$

Solution #1: First, let's follow the formula above exactly and expand upon the first row of the matrix given above, call it matrix M. Then:

$$\begin{aligned}
\det(M) &= 3 \begin{vmatrix} 4 & 0 \\ 5 & 1 \end{vmatrix} - (-1) \begin{vmatrix} 1 & 0 \\ -2 & 1 \end{vmatrix} + 2 \begin{vmatrix} 1 & 4 \\ -2 & 5 \end{vmatrix} \\
&= 3(4-0) + 1(1-0) + 2(5-(-8)) \\
&= 12 + 1 + 26 \\
&= 39
\end{aligned}$$

Solution #2: We could notice that zero in position $(2,3)$ and calculate the determinant by either expanding on the second row or the third column. Let's expand along the second row. To do this, we will use the second row of the sign matrix:

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}.$$

Then, our determinant is computed as:

$$
\begin{aligned}
\det(M) &= -(1)|M_{21}| + (4)|M_{22}| - (0)|M_{23}| \\
&= -1 \begin{vmatrix} -1 & 2 \\ 5 & 1 \end{vmatrix} + 4 \begin{vmatrix} 3 & 2 \\ -2 & 1 \end{vmatrix} - 0
\end{aligned}
$$

The advantage of the zero is one less determinant!

$$
\begin{aligned}
&= -1(-1 - 10) + 4(3 - (-4)) \\
&= -1(-11) + 4(7) \\
&= 11 + 28 \\
&= 39
\end{aligned}
$$

This process can obviously be continued to compute determinants of 4×4 matrices and larger, but along the way to compute the determinant you will pass through determinants of 3×3 matrices and then determinants of 2×2 matrices.

4.6 Eigenvalues and eigenvectors

Definition 4.3: Eigenvector

An *eigenvector* of an $n \times n$ matrix A is a nonzero vector $\vec{x}$ such that $A\vec{x} = \lambda \vec{x}$ for some scalar λ. A scalar λ is called an *eigenvalue* of A if there is a nontrivial solution $\vec{x}$ of $A\vec{x} = \lambda \vec{x}$.

So, how do we find eigenvalues? Let's look closer at the equation $A\vec{x} = \lambda \vec{x}$:

$$
\begin{aligned}
A\vec{x} &= \lambda \vec{x} \\
A\vec{x} &= \lambda I \cdot \vec{x} \\
A\vec{x} - \lambda I \cdot \vec{x} &= \vec{0} \\
(A - \lambda I)\vec{x} &= \vec{0}
\end{aligned}
$$

Then, this matrix equation has nontrivial solutions when $\det(A - \lambda I) = 0$. This equation is called the *characteristic equation*, with $\det(A - \lambda I)$ being called the *characteristic polynomial*.

Example 4.18: Eigenvalues

Determine the eigenvalues for the matrix

$$
\begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}.
$$

Solution: First form the matrix $A - \lambda I$:

$$A - \lambda I = \begin{bmatrix} 2 - \lambda & -12 \\ 1 & -5 - \lambda \end{bmatrix}.$$

Now, solve $\det(A - \lambda I) = 0$:

$$
\begin{aligned}
(2 - \lambda)(-5 - \lambda) - (1)(-12) &= 0 \\
-10 + 3\lambda + \lambda^2 + 12 &= 0 \\
\lambda^2 + 3\lambda + 2 &= 0 \\
(\lambda + 2)(\lambda + 1) &= 0 \\
\lambda &= -2, -1
\end{aligned}
$$

So, the eigenvalues for the matrix A are -2 and -1.

Now that we can find eigenvalues, how do we find the corresponding eigenvectors? Well, for a given eigenvalue λ, we just solve the matrix equation $(A - \lambda I)\vec{x} = \vec{0}$.

Example 4.19: Eigenvectors

Determine the eigenvectors for the matrix from the previous example.

Solution: Let's start by finding the eigenvectors that correspond to the eigenvalue $\lambda = -2$. So, we need to solve $(A + 2I)\vec{x} = \vec{0}$.

$$\begin{bmatrix} 4 & -12 & 0 \\ 1 & -3 & 0 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & -3 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

So, the eigenvectors that correspond to the eigenvalue $\lambda = -2$ are $\{[3y, y]^\mathsf{T} \mid y \in \mathbb{R}\}$.

Similarly, we can find the eigenvectors that correspond to $\lambda = -1$. We just need to solve $(A + I)\vec{x} = \vec{0}$.

$$\begin{bmatrix} 3 & -12 & 0 \\ 1 & -4 & 0 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & -4 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

So, the eigenvectors that correspond to the eigenvalue $\lambda = -1$ are $\{[4y, y]^\mathsf{T} \mid y \in \mathbb{R}\}$.

4.7   Exercises

1. Solve the following system by row reducing an augmented matrix to RREF:
$$
\begin{array}{rcrcrcr}
3x & + & 2y & - & z & = & 1 \\
x & + & 2y & + & 2z & = & 0 \\
2x & + & y & - & 3z & = & -1
\end{array} .
$$

2. Determine whether the following set of vectors is linearly independent or linearly dependent in $\mathbb{R}^4$:
$$
\left\{
\begin{bmatrix} 2 \\ 5 \\ 1 \\ 0 \end{bmatrix},
\begin{bmatrix} 0 \\ 1 \\ 2 \\ -1 \end{bmatrix},
\begin{bmatrix} -3 \\ 1 \\ 5 \\ 1 \end{bmatrix}
\right\} .
$$

3. Given the matrices
$$
A = \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 4 & -5 \end{bmatrix}
\quad , \quad
B = \begin{bmatrix} 4 & 1 & -1 \\ 2 & 2 & 9 \end{bmatrix}
\quad , \quad
C = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 3 & 0 \\ 0 & 1 & -1 \end{bmatrix}
$$
compute each of the following (if possible):

(a) $3A - 2B^{\mathsf{T}}$

(b) BC

(c) BA

(d) C^{-1}

4. Calculate the determinant of
$$
A = \begin{bmatrix} 2 & 1 & -1 \\ -3 & 0 & 0 \\ 4 & 1 & 5 \end{bmatrix} .
$$

5. Determine the eigenvalues, along with the corresponding eigenvectors, for
$$
M = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} .
$$

Chapter 5

Proofs Techniques

This chapter will serve as a brief review of the essential concepts needed when proving or disproving mathematical statements. Some of the concepts presented in this chapter overlap with content in the chapter on calculus.

Mathematical Statements are either true or false.

> **Example 5.1: Statements**
>
> The following are mathematical statements:
>
> 1. 3 is even.
>
> 2. $f(x)$ is differentiable for $x \in \mathbb{R}$, where $f(x) = x^2$.
>
> 3. All prime numbers are odd, for $p \geq 3$.

For clarity it is often useful to use "if-then" statements.

p	q	$p \longrightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

5.1 Proof Techniques

In this section a verity of proof techniques and some examples will be discussed.

5.2 $\epsilon - \delta$ Style Proofs

For limits, sequences, and series this style proof is popular. More information about this style proof technique is given in Chapter 2 that reviews concepts from calculus. A few examples are given here to illustrate this style of proof.

Definition 5.1: Converging Sequence

A sequence $\{a_n\}$ of real numbers converges to $L \in \mathbb{R}$ if $\forall\ \epsilon > 0, \exists N_0 \in \mathbb{N}$ such that if $n > N_0$, then $|a_n - L| < \epsilon$.

Example 5.2: Sequence Converges

Prove that $\left\{ \frac{n+5}{n+1} \right\}$ converges to 1.

Proof. Consider

$$\{s_n\} = \left\{ \frac{n+5}{n+1} \right\}_{n \geq 1}$$

Let $\epsilon > 0$. Choose $N_0 = \lceil \frac{4}{\epsilon} \rceil$. Now, suppose $n > N_0$. So, $n > \frac{4}{\epsilon}$, and $n > \frac{4}{\epsilon} - 1$. Thus, $\frac{4}{n+1} < \epsilon$. This means

$$\left| \frac{4}{n+1} \right| < \epsilon \Longrightarrow \left| \frac{n+5}{n+1} - 1 \right| < \epsilon.$$

Therefore, $\left\{ \frac{n+5}{n+1} \right\}$ converges to 1. $\square$

Definition 5.2: Limit

For $f : S \to \mathbb{R}$, with $S \subseteq \mathbb{R}$, we say

$$\lim_{x \to a} f(x) = L \text{ if } \forall \epsilon > 0, \exists \delta > 0$$

such that

$$\text{if } 0 < |x - a| < \delta, \text{ then } |f(x) - L| < \epsilon.$$

Using the limit definition prove $\lim_{x \to 3}(4x + 7) = 19$.

> **Example 5.3: Prove** $\lim_{x \to 3}(4x + 7) = 19$.
>
> *Proof.* Let $f(x) = 4x + 7$ and $a = 3$, $L = 19$, and lastly let $\epsilon > 0$.
> Set $\delta = \frac{\epsilon}{4}$. Now, let x be such that
>
> $$0 < |x - 3| < \delta.$$
>
> Then,
>
> $$|x - 3| < \frac{\epsilon}{4} \implies 4|x - 3| = |4x - 12| = |4x + 7 - 19| < \epsilon.$$
>
> Thus, the limit exists and
>
> $$\lim_{x \to 3} f(x) = 19.$$
>
> $\square$

5.2.1   Mathematical Induction

When a statement $P(n)$ is based on an index n, and the goal is to show
that $P(n)$ is true for all values of $n \in \mathbb{N}, n \geq n_0$. The idea becomes:

1. Show $P(n_0)$ is true as a base or anchor step.

2. Show that $P(k) \implies P(k + 1)$ for some $k \geq n_0$.

In order to give some examples some definitions are given.

> **Definition 5.3: Divergent Sequence**
>
> A sequence $\{a_n\}$ of real numbers is said to <u>diverge to infinity</u> if for
> all $M \in \mathbb{N}$ there exists $N_0 \in \mathbb{N}$ such that if $n > N_0$, then $a_n > M$.

> **Definition 5.4: Infinite Series**
>
> For a sequence of real numbers $\{a_n\}$, we define the infinite series to
> be $\sum_{i=1}^{\infty} a_i$.

> **Definition 5.5: Sequence of Partial Sums**
>
> For any $n \in \mathbb{N}$, we write $s_n = \sum_{i=1}^{n} a_i$. Then, we define the sequence

of partial sums as $\{s_n\}$.

> **Definition 5.6: Converges**
>
> The series $\sum_{i=1}^{\infty} a_i$ converges to $L \in \mathbb{R}$ if $\{s_n\}$ converges to L. If $\{s_n\}$ diverges to infinity, then so does the series.

With the given definitions defined we can consider the following example:

> **Example 5.4: Show $\sum_{i=1}^{\infty} \frac{1}{2^i}$ converges.**
>
> Start by considering the values of $\{s_n\}$.
>
> $$s_1 = \frac{1}{2}, s_2 = \frac{3}{4}, s_3 = \frac{7}{8}, s_4 = \frac{15}{16}$$
>
> Conjecture that
> $$s_n = \frac{2^n - 1}{2^n} \text{ for } n \geq 1.$$
>
> Thus, we are conjecturing that
>
> $$\sum_{i=1}^{n} \frac{1}{2^i} = \frac{2^n - 1}{2^n} \text{ for } n \geq 1.$$
>
> Lets use induction to prove this claim:
>
> *Proof.* Let $n \in \mathbb{N}$, with $n \geq 1$.
> **Base or Anchor Step:**
>
> $$\sum_{i=1}^{1} \frac{1}{2^i} = \frac{1}{2} \text{ and } \frac{2^1 - 1}{2^1} = \frac{1}{2}.$$
>
> As $\frac{1}{2} = \frac{1}{2}$, the result holds when $n = 1$.
> **Inductive Step:** Assume true for some $k \geq 1$. So,
>
> $$\sum_{i=1}^{k} \frac{1}{2^i} = \frac{2^k - 1}{2^k}.$$

Then,

$$\sum_{i=1}^{k+1} \frac{1}{2^i} = \sum_{i=1}^{k} \frac{1}{2^i} + \left(\frac{1}{2^{k+1}} \right)$$

$$= \frac{2^k - 1}{2^k} + \frac{1}{2^{k+1}} = \frac{2^{k+1} - 2 + 1}{2^{k+1}} = \frac{2^{k+1} - 1}{2^{k+1}}$$

Therefore by the principal of mathematical induction,

$$\sum_{i=1}^{n} \frac{1}{2^i} = \frac{2^n - 1}{2^n}.$$

$\square$

Next show $\left\{ \frac{2^n - 1}{2^n} \right\}$ converges to 1. Then, conclude the series also converges to 1.

5.3 Proof by Contradiction

When a direct proof doesn't work this method can be used to show something does not have a property or that something does not exist.

The idea here is to prove

$$p \implies q,$$

suppose p and $\tilde{q}$. Then show this leads to a contradiction.

Example 5.5: Infinitely Many Primes

Prove that there are infinitely many prime numbers. A sketch of the proof is given here.

Proof. Assume, to contradict, there are finitely many prime numbers. Lets assume that there are a prime numbers. So a list of all prime numbers would be:

$$P_1, P_2, P_3, P_4, \ldots, P_{a-1}, P_a.$$

Then, set

$$M = (P_1 \cdot P_2 \cdot P_3 \cdots P_a) + 1.$$

$M > 1$, so there exists a prime in the list such that $P_k | M$. Thus,

$$P_k \Big| [(P_1 P_2 \cdots P_k \cdots P_a) + 1],$$

and

$$P_k \Big| (M - P_1 \cdot P_2 \cdots P_a) \implies P_k | 1 \implies P_k = 1.$$

Here P_k is prime and $P_k = 1$, a contradiction. $\square$

5.4   Direct Proof

This method should be implemented as often as possible when looking to show that $p \implies q$. The idea is to show $p \implies q$ is true, suppose p is true and then show that q is a logical consequence of p.

5.5   Proof by Contrapositive

This proof technique is used in similar places to proof by contradiction. It is employed when q is less complicated than p to work with. The idea is to show $p \implies q$, suppose that $\tilde{q}$ and then show $\tilde{p}$. Thus, you are proving:

$$\tilde{q} \implies \tilde{p}.$$

Example 5.6: For $x \in \mathbb{Z}, 3x - 7$ is even iff x is odd.

Proof. Let x be an integer.

$\Leftarrow$ Suppose x is odd. Then,

$$\exists\, n \in \mathbb{Z} \text{ such that} x = 2n + 1.$$

$$3x - 7 = 3(2n + 1) - 7 = 6n - 4 = 2(3n - 2)$$

and $3n - 2 \in \mathbb{Z}$. Thus, $3x - 7$ is even.

$\Rightarrow$ Showing the contrapositive. Assume x is not odd, which means x is even. Then, $\exists\, m \in \mathbb{Z}$ such that $x = 2m$. Now

$$3x - 7 = 3(2m) - 7 = 6m - 7 = 6m - 8 + 1 = 2(3m - 4) + 1$$

where $3m - 4 \in \mathbb{Z}$. Thus, $3x - 7$ is odd, as desired.

Therefore $3x - 7$ is even if and only if x is odd.   □

5.6   Subset and Set Equality Proofs

When considering specific sets or set identities this method of proof is employed. The idea is given two sets A and B:

1. To show $A \subseteq B$, show arbitrary $x \in A$ is also in B.

$$\text{If } x \in A, \text{ then } x \in B.$$

2. To show $A = B$, show $A \subseteq B$, and $B \subseteq A$.

$$x \in A \text{ iff } x \in B.$$

5.7   Existence and/or Uniqueness Proofs

For a uniqueness proof the idea becomes to let x_1, x_2 have a given property. Then show $x_1 = x_2$.

5.8   Function Proofs

Function proofs are used to show properties like a function is one-to-on, (1-1), or onto.

5.9   (If and only if): $p \longleftrightarrow q$

The main idea when examining $p \longleftrightarrow q$ is to prove both: $p \implies q$ and $q \implies p$.

5.10   Exercises

1. Given $a_n = \left\{ \frac{1}{n} \right\}$, find and prove a closed form for s_n. Then show $\{s_n\}$ diverges.

2. Given $f(x) = 2x - 9$ use a $\epsilon - \delta$ proof to show the $\lim_{x \to 5} f(x) = 1$.

3. Let $n \in \mathbb{Z}$. Prove that $2n^2 + n$ is odd if and only if $\cos(\frac{n\pi}{2})$ is even.

4. Use the three parts described below to prove:

$$\sum_{k=1}^{\infty} \frac{1}{(3k - 2)(3k + 1)} \text{ converges to } \frac{1}{3}.$$

(a) Let $\{s_n\}$ be the sequence of partial sums for $n \geq 1$. Compute the first several values of the sequence and then conjecture a closed form formula for a general term of $\{s_n\}$.

(b) Use induction to prove your conjecture from part (a).

(c) Using the result of part (b) as a lemma, prove that the sequence of partial sums converges to $\frac{1}{3}$, and use this to conclude that the original series converges to $\frac{1}{3}$ as well.

Chapter 6

Statistics Review

This chapter will serve as a brief review of the essential concepts encountered in statistics course.

6.1 Probability

Definition 6.1: probability space

The triple (Ω, β, P) is called a probability space where

- Ω is the sample space.

- β is the set consisting of all events we can assign as outcomes. β is also called an event space.

- P is a probability measure that assigns probabilities to the events in β.

The probability measure must satisfy the conditions:

- For any event $A \in \beta$, $P(A) \geq 0$.

- $P(\Omega) = 1$.

- Let $A_1, A_2, \ldots$ be events such that $A_i \cap A_j = \varnothing$ for all $i \neq j$ then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Theorem 6.1: Probability Theorems

1. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

2. **Complement:** $P(A) = 1 - P(A^c)$ or $P(A) = 1 - P(\bar{A})$.

3. $P(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$.

Definition 6.2

he conditional probability that event A occurs given that B occurs denoted as $P(A|B)$ is defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{6.1}$$

Read $P(A|B)$ as "the probability of A given B."

Rearranging (6.1) results in

$$P(A \cap B) = P(A|B)P(B).$$

Definition 6.3: Independent Events

Events A and B are independent if and only if either of the following occur:

1. $P(A|B) = P(A)$

2. $P(B|A) = P(B)$

Theorem 6.2: Independent Events

Events A and B are independent if and only if $P(A \cap B) = P(A)P(B)$.

Definition 6.4: Mutually Independent

Events $A_1, A_2, \ldots, A_k$ are said to be mutually independent if and

only if for any $I \subseteq \{1, 2, \ldots, k\}$,

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i).$$

6.2 Random Variables and Probability Distributions

Definition 6.5: Probability Mass Function

A probability mass function (pmf) is a function that assigns each element of the sample space a real number.

Capital letters such as X, Y, and Z are used to denote a random variable. Random variables are typically classified as being either discrete or continuous. It is also possible to considered a mixed random variable.

Definition 6.6: Expected Value

The mean (expected value) and variance of a discrete random variable X are defined as

$$\mu = \sum_x xP(X = x)$$

and

$$\sigma^2 = V(X) = \sum_x (x - \mu)^2 P(X = x) = \sum_x x^2 P(X = x) - \mu^2.$$

Definition 6.7: Mean Discrete Random Variable

Let $g(\cdot)$ be a function. Then $E(g(X)) = \sum_x g(x)P(X = x)$.

The variance of a random variable is given by:

$$V(X) = E(X^2) - \mu^2.$$

Definition 6.8: Binomial Experiment

n experiment is said to be a binomial experiment if

1. The experiment is performed n times and n is fixed. Each repetition of the experiment is called a trial.

2. The trials are independent.

3. Each trial results in one of two outcomes. One outcome is called a success and the other a failure.

4. For each trial, P(success)=p and P(failure)= $(1 - p)$.

In a binomial experiment the random variable of interest is $X =$ number of successes out of the n trials.

Definition 6.9: Discrete Random Variable

A discrete random variable X is said to have a binomial distribution with parameters n and p ($X \sim \text{Binomial}(n, p)$) if

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

for $x = 0, 1, 2, \ldots, n$ and $p \in [0, 1]$.

In the case of a discrete random variable the expected value $E(X) = np$ and the variance is defined by $V(X) = np(1 - p)$.

Definition 6.10: Poission Distribution

A discrete random variable X is said to have a Poisson distribution with parameter λ ($X \sim \text{Poisson}(\lambda)$) if

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

for $x = 0, 1, 2, \ldots$ and $\lambda > 0$.

Example 6.1: Poisson Random Variable

Find the mean of the Poisson random variable.

$$
\begin{aligned}
E(X) &= \sum_{x=1}^{\infty} x P(X = x) \\
&= \sum_{x=1}^{\infty} x \frac{e^{-\lambda} \lambda^x}{x!} \\
&= e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!} \\
&= e^{-\lambda} \left(\lambda e^{\lambda} \right) \\
&= \lambda
\end{aligned}
$$

Definition 6.11: Probability Density Function

The probability density function (pdf) of a continuous random variable X, denoted as $f(x)$ is a function that must satisfy the following:

1. For all x, $f(x) \geq 0$.

2. The area under $f(x)$ is 1. ($\int_{-\infty}^{\infty} f(x) dx = 1$.)

3. $P(a \leq X \leq b)$ is area under the curve between a and b.
   $P(a \leq X \leq b) = \int_a^b f(x) dx$.

Definition 6.12: Expected Value

Let X be a continuous random variable with probability density function $f(x)$. The mean (expected value) and variance of X are defined as

$$
\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx
$$

and

$$
\sigma^2 = V(X) = E((X - \mu)^2) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx
$$
$$
= \int_{-\infty}^{\infty} x^2 f(x) dx - \mu^2 = E(X^2) - \mu^2.
$$

If a and b are non-zero constants then $E(aX + b) = aE(X) + b$ and $V(aX + b) = a^2 V(X)$.

Example 6.2: Create New Random Variables

Suppose $E(X) = -5$, $V(X) = 3$, and $Y = 2X - 5$. Find the mean, variance, and standard deviation of Y.

$$E(Y) = E(2X - 5) = 2E(X) - 5 = 2(-5) - 5 = -15$$

$$V(Y) = V(2X - 5) = 2^2 V(X) = 4(3) = 12$$

$$\implies \text{ the standard deviation of } Y \text{ is } \sqrt{12}.$$

Definition 6.13: Joint Probability Mass Function

Let X and Y be two random variables.

1. If X and Y are discrete random variables then the joint probability mass function is $p(x, y) = P(X = x, Y = y)$.

2. If X and Y are continuous random variables then the joint probability density function is a function $f(x, y)$ that satisfies

    (a) $f(x, y) \geq 0$ for all $(x, y) \in \mathbb{R}^2$ and
    (b) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$.

Theorem 6.3: Expected Value of Bivariate

Let X and Y be continuous random variables and $g(\cdot, \cdot)$ be a bivariate function then

$$E(g(X, Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy.$$

Definition 6.14: Covariance

Let X and Y be random variables with means μ_X and μ_Y respectively. The covariance between X and Y is

$$\sigma_{XY} = \text{Cov}(X, Y) = E((X - \mu_x)(Y - \mu_Y)).$$

When considering the covariance between two random variables X and Y note that:

1. $-\infty < \text{Cov}(X, Y) < \infty$.

2. The covariance measures the strength of the associating relationship between two random variables.

3. $\text{Cov}(X, X) = V(X)$.

4. The correlation between random variables X and Y is

$$\rho_{XY} = \text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

It can be shown that $-1 \le \rho_{XY} \le 1$.

5. If the random variables X and Y are independent $\text{Cov}(X, Y) = 0$.

6. If $a_1, \ldots, a_n$ are non-zero constants and $X_1, \ldots, X_n$ are random variables then

$$V\left(\sum_{i=1}^{n} a_i X_i\right) = \underbrace{\sum_{i=1}^{n} a_i^2 V(X_i)}_{\substack{\text{Sum of the} \\ \text{individual} \\ \text{variances.}}} + \underbrace{2\sum_{i<j} a_i a_j \text{Cov}(X_i, X_j)}_{\substack{\text{Takes into account the} \\ \text{dependency between the} \\ \text{random variables.}}}.$$

7. If $X_1, \ldots, X_n$ are independent then

$$V\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 V(X_i)$$

Definition 6.15: Conditional Density Function

Let X and Y be continuous random variables. The conditional density function of X given Y denoted as $f(x|y)$ is

$$f(x|y) = \frac{f(x, y)}{f(y)}.$$

A similar definition exists for discrete random variables, and the bivariate case for random variables can be extended to a multivariate case.

# 6.3 Normal Distribution

> **Definition 6.16: Normal Distribution**
>
> A random variable X is said to have a normal distribution if its probability density function is
>
> $$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\frac{-(x - \mu)^2}{2\sigma^2}\right\}$$
>
> where $-\infty < x < \infty$, $-\infty < \mu < \infty$ and $\sigma^2 \geq 0$.

The following are key features of the normal distribution.

1. Notation $N(\mu, \sigma^2)$ is typically used to denote a normal distribution with mean μ, and variance σ^2.

2. $E(X) = \mu$ and $V(X) = \sigma^2$.

3. The graph of the probability density function is
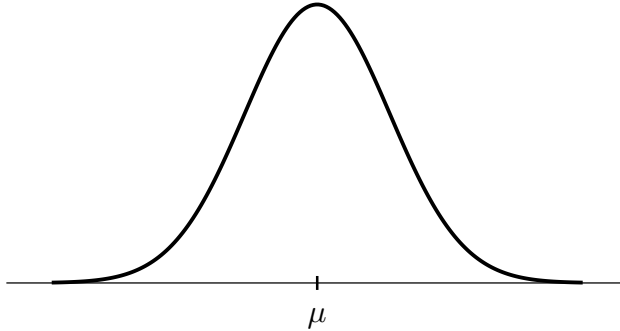
4. $Z \sim N(0, 1)$ is called the standard normal distribution.

5. If $X \sim N(\mu, \sigma^2)$ then

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1).$$

6. If $X_1, \ldots, X_n$ are independent $N(\mu, \sigma^2)$ random variables then $\bar{X} \sim N(n\mu, n\sigma^2)$.

7. If $X_1, \ldots, X_n$ are independent random variables with $E(X^2) < \infty$ then by the central limit theorem,

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \text{ converges in distribution to } Z \sim N(0, 1)$$

as $n \to \infty$.

8. If X and Y are normally distributed random variables with $\mathrm{Cov}(X, Y) = 0$ then X and Y are independent. **Note: In general this is not true.**

6.4 Chi-Squared Distribution

> **Definition 6.17: Chi-Squared Distribution**
>
> Let $Z_1, \ldots, Z_v$ be independent standard normal random variables then $Q = \sum_{i=1}^{v} Z_i^2$ has a chi-squared (or χ-squared) distribution with v degrees of freedom.

Important properties of the chi-squared distribution include:

1. Typical notation includes $Q \sim \chi^2(v)$ or $Q \sim \chi_v^2$.

2. The graph of the probability density function is
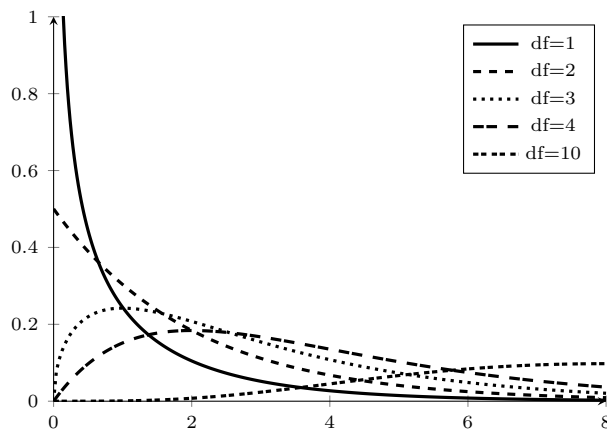
3. $E(Q) = v$ and $V(Q) = 2v$.

4. If $X_1, \ldots, X_n$ are independent $N(\mu, \sigma^2)$ random variables then

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$

6.5 Student's t-Distribution

Definition 6.18: Student's t-Distribution

Let $Z \sim N(0,1)$ and $Q \sim \chi^2(v)$ be **independent** random variables. Then
$$t = \frac{Z}{\sqrt{Q/v}}$$
has a Student's t-distribution with v degrees of freedom.

The key points of Student's t-distribution include:

1. Notation: $t \sim T(v)$

2. The graph of the probability density function is

$$0$$
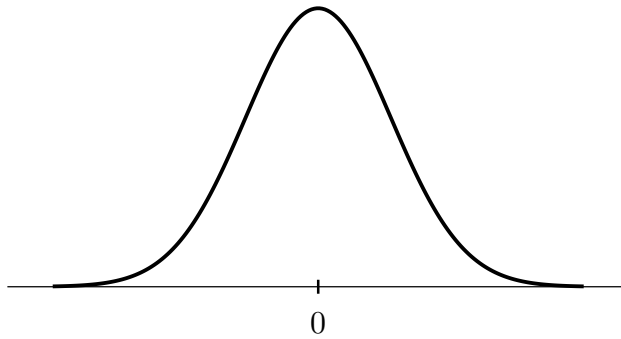
3. If $X_1, \ldots, X_n$ are independent $N(\mu, \sigma^2)$ random variables then

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} \sim Z \text{ or } N(0,1).$$

6.6 F-Distribution

We now look at Snedecor's F-distribution quickly.

Definition 6.19: F-Distribution

Let $Q_1 \sim \chi^2(v_1)$ and $Q_2 \sim \chi^2(v_2)$ be **independent** random variables. Then
$$F = \frac{Q_1/v_1}{Q_2/v_2}$$
has an F-distribution with v_1 numerator degrees of freedom and v_2 denominator degrees of freedom.

Facts:

1. Notation: $F(v_1, v_2)$.

2. The graph of the probability density function is given below.
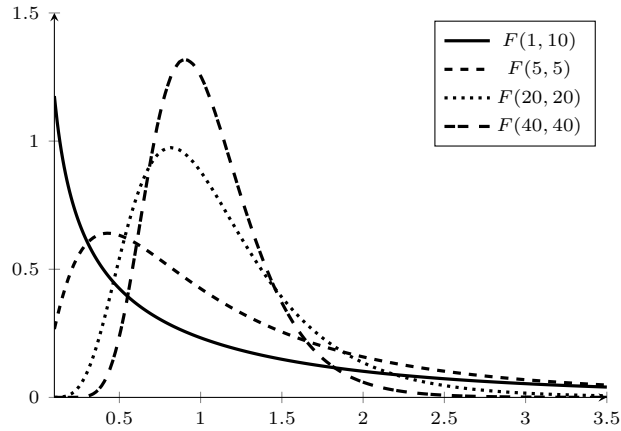
3. If $F \sim F(v_1, v_2)$ then $1/F \sim F(v_2, v_1)$.

4. If $t \sim t(v)$ then $t^2 \sim F(1, v)$.

   Other important distributions exist including the negative binomial, hypergeometric, uniform, exponential, beta, gamma, Weibull, and Cauchy distributions.

6.7   Methods of Estimation

Definition 6.20: Point Estimate

A point estimate is a single number used to estimate a population parameter.

Definition 6.21: Confidence Interval

A confidence interval is used to express the precision and uncertainty associated with a particular sampling method.

6.7.1 $(1 - \alpha) \times 100\%$ Confidence Interval Notation

We call the interval a $(1 - \alpha) \times 100\%$ confidence interval. We will need to determine α when doing the problems.

Example 6.3: 95% Confidence Interval

Find α for a 95% confidence interval:

$$\implies \alpha = 0.05$$

$(1 - \alpha) \times 100\%$ Confidence Interval for μ

We begin with the following assumptions:

1. Assume that a sample of size n is taken from a normally distributed population.

2. The population standard deviation σ is assumed to be known.

Based upon the above assumption, we can show that

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

This fact can be used to show that the formula for the $(1 - \alpha) \times 100\%$ confidence interval for μ is

$$\left(\bar{x} - t_{\alpha/2}(n-1)(s/\sqrt{n}), \ \bar{x} + t_{\alpha/2}(n-1)(s/\sqrt{n})\right) = \bar{x} \pm t_{\alpha/2}(n-1)(s/\sqrt{n}).$$

Confidence intervals can be constructed for all types of parameters including μ_D, $\mu_1 - \mu_2$, p, and $p_1 - p_2$.

6.8   Hypothesis Testing

Court Case Analogy:

- At the start of a trial, a person is assumed to be innocent until proven guilty.

- The evidence must be beyond a reasonable doubt to get a guilty verdict.

- We would rather let a guilty person go free then send an innocent person to jail.

- If there is not enough evidence to get a guilty verdict then we say that the person is not guilty. We never say that the person is guilty.

> **Definition 6.22: Null Hypothesis**
>
> The null hypothesis is the hypothesis to be tested. We assume that this hypothesis is true. ("An individual is assumed to be innocent.") We denote the null hypothesis as H_0.

> **Definition 6.23**
>
> he alternative hypothesis is the statement that is true if the null hypothesis is false. ("An Individual is Guilty.") We denote the alternative hypothesis as H_a.

The **alternative hypothesis** is also called the **Researcher's Hypothesis**.

Decisions:

In order to evaluate a hypothesis test:

- We take a random sample from the population.

- If the sample indicates there is enough evidence to believe that the null hypothesis is false we **reject** the null hypothesis and **assert** the alternative hypothesis.

- If the sample does not indicate that the null hypothesis is false we **fail to reject** the null hypothesis and indicate that there is **not enough evidence** to **assert** the alternative hypothesis.

WE NEVER ACCEPT THE NULL HYPOTHESIS OR SAY THAT THE NULL HYPOTHESIS IS TRUE.
(A PERSON IS PRONOUNCED NOT GUILTY. WE DO NOT SAY THEY ARE INNOCENT).

6.9 Errors in Hypothesis Testing:

Considering the types of errors that can occur.

Decision	Reality	
	H_0 is True	H_0 is False
Reject H_0	Type I Error	No Error
Fail to Reject H_0	No Error	Type II Error

Type I Error: **We put an innocent person in jail**.
Type II Error: **We let a guilty person go free**.
    We want to control both Type I and Type II errors. If we **lower** one
type of error we **increase** the other type of error.

Definition 6.24: Significance Level

The probability of making a **Type I Error** is called the **signifi-
cance level**. It is denoted by α.

$$P[\text{Type I Error}] = \alpha$$

- We want to minimize Type I errors therefore we choose small values
  for α. Typical values are .01, .05, or .10.

- If we reject the null hypothesis and assert the alternative hypothesis
  we say the test

  results are **statistically significant** at the α level.

- If we fail to reject the null hypothesis we say that the test results
  are **not statistically significant** at the α level.

$$P[\text{Type II Error}] = \beta$$

Definition 6.25: Power of Hypothesis Test

The probability of rejecting the null hypothesis given that it is false
is called the **power** of a hypothesis test.

$$\text{Power} = 1 - \beta = P(\text{Reject } H_0 | H_0 \text{ is false })$$

6.10   Hypothesis Testing for μ

To start:

Assumptions:

1. Assume that a random sample is taken from a normally distributed
   population.

2. Also assume that μ is unknown.

Hypotheses:

Three cases can be tested:

$$H_0 : \mu = \mu_0 \qquad H_0 : \mu = \mu_0 \qquad H_0 : \mu = \mu_0$$
$$H_a : \mu < \mu_0 \qquad H_a : \mu > \mu_0 \qquad H_a : \mu \neq \mu_0$$

Test Statistic:

If the null hypothesis is true then

$$t^* = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t(n-1).$$

6.11 Exercises

1. X is a discrete random variable whose probability distribution is given below.

x	0	1	2	3
$P(X = x)$	0.2	0.3	0.4	0.1

   Find the mean, variance, and standard deviation of X.

2. Let X be a random variable whose pdf is $f(x) = 3e^{-3x}$. Find $P(X \geq x)$ and $E(X)$.

3. Suppose $E(X) = 30$, $V(X) = 10$, and $Y = 100 - 8X$. Find the expected value, variance, and standard deviation of Y.

4. Let $Y_i = \beta_0 + \beta_1 x_i + \mathcal{E}_i$ where $\mathcal{E}_i$s are independent with $E(\mathcal{E}_i) = 0$ and $V(\mathcal{E}_i) = \sigma^2$. Find the mean and variance of Y_i for a given value of x_i.

5. A sample of 15 U.S. commercial aircraft yielded a sample mean age of 14.25 years and a standard deviation of 9.35 years. Construct a 90% confidence interval for the true mean age of all U.S. commercial aircraft and give an interpretation.

6. Fat contents (in percentage) for hot dogs is approximately normally distributed. A sample of 10 randomly selected hot dogs was taken and the fat content was determined for each hot dog. Construct a 99% confidence interval for the true mean fat percentage of hot dogs.

   25.2   21.3   22.8   17   29.8   21   25.5   16   20.9   19.5

7. You are testing $H_0 : \mu = 45$ versus $H_a : \mu > 45$ using a .10 significance level. After taking a random sample of size 63 you calculate your test statistic and obtain $t^* = 2.17$. Give the p-value and decision.

8. A soda company distributes soda in cans labeled 12 oz. The Bureau of Weights and Measures randomly selects 36 cans, measures their contents, and obtains a sample mean of 11.82 oz and a sample standard deviation of .38 oz. At the .01 significance level is there enough evidence to conclude that the soda company is cheating consumers.

Chapter 7

Scientific Computing Resources

The need to use a programming language or do numerical simulation when solving mathematical problems is constantly increasing. This chapter has a few places where you can look to find the basics of programming. The goal here is to be a starting point for the reader. The subsections give a starting point for the reader to find some functional web-based resources that will get them coding quickly.

7.1   Anaconda

A good one stop shop for students to get a start with a lot of the tools that will be used in your different course work is Anaconda. It is a full product bundling for different operating systems including both Windows, Linux, and Mac OS. A webpage that would get you started is:

```
https://www.anaconda.com/products/individual
```

The base open source tool kit is free, and includes installers for both integrated development environments like jupyter, and spyder, but also the programming languages that will be used in your course work like python, and R as well as their associated packages like pandas and numpy.

There is an Anaconda Navigator desktop GUI that allows you to launch applications and manage packages and environments with out the need to install them on you machines command line interface. The page also includes training materials for data science and your programming needs.
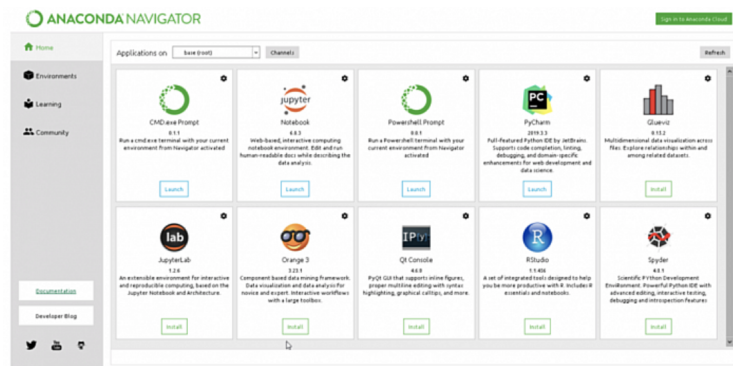
7.2   Collaborative Computing in The Cloud

Web based platforms for doing computational science have are starting to really take off. Most of these resources have a place where users can register a free account and then start using the tools for their own problems. This has the added benefit of novice users not needing to install lots of software out of the gate.

One of the best collections of tools can be found on

$$\text{https://cocalc.com/}$$

COCALC, or Collaborative Calculation and Data Science has a collection of tools all in one place for students to use. Included in the online set of tools:

1. Jupyter Notebooks

2. A LaTeX Editor

3. An Online Linux Environment

4. Octave - A mathematical programming syntax largely compatible with MATLAB.

5. R - Statistical computing environment.

6. Python - remote environment allowing for scripting and interactive notebooks, and graphical application development.

7.3   Integrated Development Environments

There are lots of integrated development environments (IDE) that can be used for scientific computing. A quick web search will produce more

than you can easily invest your time in. Each has its own tutorials that can assist a user with getting it up and running, and some have nice web-based area's where you can test drive them online. To be concise in this document I will only point out two that have been fairly easy for students to get up and running:

7.3.1 Spyder

The Scientific Python Development Environment (Spyder) has a full set of tools for interactive python development.

$$https://www.spyder-ide.org/$$

The resource is free and open source and dose some very nice data exploration, and visualization for scientific computing.

7.3.2 Jupyter

Jupyter is a web application that can be used to create and share scientific documents. The underlying compute kernel can be set to many different programming languages including: sage, R, python, octave, and C++. You can easily test out Jupyter's notebook and lab environments online at:

```
https://jupyter.org/
```

Using the online web interface here allows a user a nice starting point and the chance to see some well written tutorials in an open-source notebook setting. The reader is encouraged to explore at the Lorenz System that can be found at:

```
https://jupyter.org/try
```

Note some great tutorials for Jupyter are available. Look at:

```
https://www.dataquest.io/blog/jupyter-notebook-tutorial/
```

for a jumping off point from the people at Dataquest.

7.4 An Hour of Python

Trinket has come up with some nice tutorials for basic python computation. These can be found at:

```
https://hourofpython.com/
```

Their are lots of basic coding examples to get a user started. This includes coding with python Turtles, and interactive code blocks that showcase how procedural programming is accomplished.

7.5   Trinket

For a user that wants to be able to have a coding environment on any web-based document trinket offers embedded coding frameworks.

$$https://trinket.io/$$

Using the framework from the Trinket site you can run, write, and share code on any device. These code spinets can be placed into a course or learning management system for students to share interactively their work.

7.6   Shodor

For a student looking to learn how to code, Shodor is a National resource for computational science education.

$$www.shodor.org$$

There is a full set of tools and tutorials for scientific computing on their site.

7.7   Exercises

1. Explore Lorenz System that can be found at:

   ```
   https://jupyter.org/try
   ```

   Be sure to use the interactive sliders that show how model parameters effect the system.

2. Use one of the web-based resources above to create Jupyter notebook that will assist you in visualizing the output of a mathematical function.

3. Share an interactive notebook on CoCalcs website with another person and see how you can both collaborate on a scientific work simultaneously.

Chapter 8

Data Sets and R Software Tool

The R software package is often used when looking to model or gain understanding using large data sets. There are some nice tutorials for the R available for free on line. One of the places you can look for nice tutorials in many programming languages is software carpentry. A starting point tutorial for R can be found at:

https://swcarpentry.github.io/r-novice-inflammation/

The following are some notes on R that are still under construction.

This chapter will serve as a brief introduction to looking at large data sets and manipulating and exploring these data sets using the R software tool.

8.1   Tidy Data

The typical workflow for data analysis looks something like:
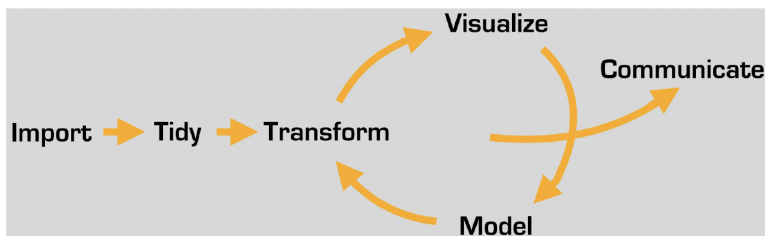
Notice the feedback between the visualization, Model, and Transforming the Data. In order to easily work with data in software tools like R it is

best to ensue that the data is tidy. Specifically it is useful to have each variable be a single column in the data set and each observation has its own row.

Each **variable** must have its own **column**   Each **observation** must have its own **row**   Each **value** must have its own **cell**
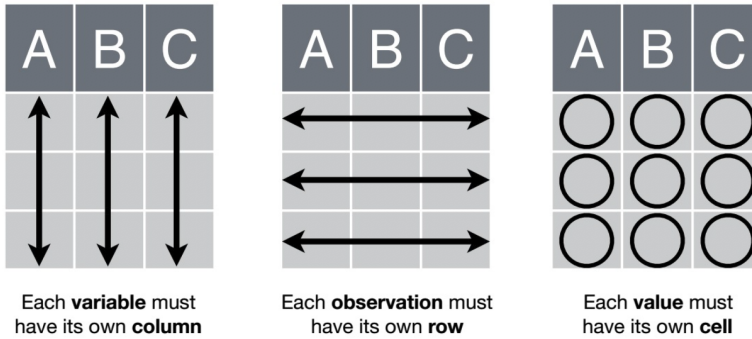
Once your data is organized you are ready to start manipulating and analyzing it using R.

8.2   Running R

The R software can be run using R studio or through jupyter style notebooks. Look at:

$$\text{https://rstudio.com}$$

and

$$\text{https://jupyter.org}$$

for more information about installing these tools locally on your own machine.

In either environment you can run R commands that will allow you to work with and analyze your data.

Example 8.1: Installing a Library

Use the following command to install the "tidyverse" library.

```
install.packages("tidyverse")
```

Then you can load the library using:

```
library(tidyverse)
```

Most operations will be done using function calls.

> **Example 8.2: Function Calls**
>
> Simple operations can be done using standard function calls:
>
> ```
> sqrt(4)
> ```
>
> Nested function calls are allowed as well using a jupyter notebook cell try:
>
> ```
> ans1 = c(7,42,1,25)
> ans2 = log(ans1)
> ans3 = round(ans2,2)
> ans4 = diff(ans3)
> ans5 = diff(round(log(c(7,42,1,25)), 2))
> print(ans1)
> print(ans2)
> print(ans3)
> print(ans4)
> print(ans5)
> ```
>
> Note that the four elements are combind, then the natural log is taken element by element. The answer is rounded to 2 decimal places, and then the difference between consecutive elements is found. The returned output is given as:
>
> ```
> [1]  7 42  1 25
> [1] 1.945910 3.737670 0.000000 3.218876
> [1] 1.95 3.74 0.00 3.22
> [1]  1.79 -3.74  3.22
> [1]  1.79 -3.74  3.22
> ```

You can define functions to assist your work flow.

> **Example 8.3: Define Functions**
>
> The following defines the subtraction of two items.
>
> ```
> my_func = function(x,y){x - y}
> ```
>
> Then it can be run by calling the function with variables or values.

8.3   Exercises

1. Try defining a function in R. Then call this function with several values or predefined variables.

Index