# Contents


**1 Taylor's Theorem** **2**
- 1.1 Analytic Functions and Taylor Series . . . 2
- 1.2 Taylor's Theorem with Remainder . . . 5
- 1.3 The Binomial Theorem . . . 6
  - 1.3.1 Properties of Binomial Coefficients . . . 8
  - 1.3.2 Negative and Non-Integer Exponents . . . 8
- 1.4 Multidimensional Taylor Series . . . 8
  - 1.4.1 Error in the Taylor Expansion . . . 10
- 1.5 Extrema of Multivariate Functions . . . 10
- 1.6 Lagrange Multipliers . . . 11

**2 First Order Differential Equations** **13**
- 2.1 Introduction . . . 13
- 2.2 Separation of Variables . . . 13
- 2.3 Differential Forms . . . 14
- 2.4 Exact Equations . . . 15
- 2.5 Integration Factors . . . 16
- 2.6 Variation of Parameters . . . 17
- 2.7 Existence and Uniqueness of Solutions for First-Order ODEs . . . 18

**3 Second Order Differential Equations** **20**
- 3.1 Introduction . . . 20
- 3.2 Constant Coefficients . . . 20
- 3.3 Complex Roots . . . 21
- 3.4 Method of Reduction of Order . . . 23
- 3.5 Variation of Parameters . . . 24
- 3.6 Method of Undetermined Coefficients . . . 26
- 3.7 Existence and Uniqueness . . . 29

**4 Eigenvalues and Eigenvectors** **30**
- 4.1 Definition of Eigenvectors and Eigenvalues . . . 30
- 4.2 Computing Eigenvalues and Eigenvectors . . . 31
- 4.3 Diagonalization . . . 34
- 4.4 Spectral Theorem . . . 36
- 4.5 Generalized Eigenvectors . . . 36
- 4.6 Matrix Exponentials . . . 37
- 4.7 The Fundamental Solution of a Matrix . . . 38

# Chapter 1

# Taylor's Theorem

## 1.1 Analytic Functions and Taylor Series

In this section we look to develop a method to represent functions as series. An important application of such is the use of series as solutions to differential equations.

**Definition 1.1 (Power Series)**

Let $x_0 \in \mathbb{R}$. A **power series** cenetered at $x_0$ is in the form

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n.$$

**Remark.** *The convention used for $0^0$ when $x = x_0$ here is $0^0 = 1$. You can imagine that it is the limit of $(x - x_0)^0$ as $x \to x_0$.*

**Remark.** *The series might not converge. In fact, the root test for convergence gives an exact criteria for convergence.*

**Definition 1.2 (Radius of Convergence)**

Let $f(x)$ be a power series centered at $x_0$. The **radius of convergence** is the value $R$ such that

$$\begin{cases} f(x) \text{ converges}, & \text{if } |x - x_0| < R, \\ f(x) \text{ diverges}, & \text{if } |x - x_0| > R. \end{cases}$$

**Definition 1.3 (Analytic Functions)**

Let $\Omega \subseteq \mathbb{R}$ be open, and $f : \Omega \to \mathbb{R}$. We say that $f$ is **analytic** at $x_0$ if there exists $\epsilon > 0$, and a power series representation $p_{x_0}(x) = \sum_n a_n (x - x_0)^n$ such that $p_{x_0}(x)$ converges to $f(x)$ in $B(x_0, \epsilon)$. We say that $f$ is analytic on (a,b) if $f$ is analytic at every point in $(a, b)$.

**Proposition 1.4**

The set of points on which $f$ is analytic form an open set.

Being analytic is one of the strictest properties for a function. Analytic functions are infinitely differentiable (smooth).

**Theorem 1.5 (Analytic Functions are Smooth)**

Suppose $\sum a_n x^n$ is a power series representation for some function $f$ with a radius of convergence $R > 0$. Then $f$ is infinitely differentiable on $(-R, R)$.

The proof requires some analysis knowledge out of scope of the course. The hardest part is to show that you can differentiate under the summation, i.e.

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

Assuming this, we can get power series representations for $f'(x)$ and so on.

$$\begin{array}{rcccccc} f(x) = & a_0 & +a_1(x-x_0) & +a_2(x-x_0)^2 & +a_3(x-x_0)^3 & +a_4(x-x_0)^4 & +\dots \\ f'(x) = & & a_1 & +2a_2(x-x_0) & +3a_3(x-x_0)^2 & +4a_4(x-x_0)^3 & +\dots \\ f''(x) = & & & 2a_2 & +6a_3(x-x_0) & +12a_4(x-x_0)^2 & +\dots \end{array}$$

Importantly, these have the same radius of convergence as the original function (a root test can confirm this), so we get a closed form for the $k$-th derivative of $f$:

$$f^{(k)}(x) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} a_n (x-x_0)^{n-k}$$

**Theorem 1.6 (Uniqueness of Power Series)**

Let $f$ be analytic at $x_0$. Then its power series representation $\sum_n a_n(x-x_0)^n$ is unique with coefficients

$$a_k = \frac{f^{(k)}(x_0)}{k!}.$$

*Proof.* We take the general form of the $k$-th derivative, and evaluate it at $x = x_0$. This gives us

$$f^{(k)}(x_0) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} a_n (x_0-x_0)^{n-k}.$$

On the right hand side, all the terms with $n > k$ will evaluate to 0. The term with $n = k$ evaluates to $k! a_k$. This means

$$f^{(k)}(x_0) = k! a_k \implies a_k = \frac{f^{(k)}(x_0)}{k!}.$$

Therefore, if a power series exists, it must be in the form

$$\sum_n \frac{f^{(n)}(x_0)}{n!}(x-x_0).$$

$\square$

**Definition 1.7 (Taylor Series)**

The **Taylor expansion** of $f$ centered at $x_o$ is given by

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

A function locally equals to its Taylor series (about some point) if and only if it is analytic. We will see that if a function is analytic, it is equal to the Taylor series about a point everywhere the series converges.

**Theorem 1.8 (Uniqueness of Analytic Functions)**

Let $f, g : (a, b) \to \mathbb{R}$ be analytic. Suppose $f(x) = g(x)$ on some small ball $B(x_0, \epsilon)$. Then $f = g$ everywhere on $(a, b)$.

Let's consider the function $h(x) = f(x) - g(x)$. This is analytic, as we can just take the difference of the respective coefficients in the power series for $f$ and $g$. We have $h(x) = 0$ on $B(x_0, \epsilon)$. Our task now is to show that $h(x) = 0$ everywhere, so $f = g$ everywhere.

The idea now is to start with the power series about $x = x_0$

$$\sum_n 0(x - x_0)^n.$$

We can 'slide' this $x_0$ across a little bit to $x_1 \in B(x_0, \epsilon)$ to get

$$\sum_n 0(x + x_1 - x_0 - x_1)^n = \sum_n 0(x - x_1)^n$$

after binomial expansion of the terms $((x - x_1) + (x_1 - x_0))^n$. Since $f$ is analytic at $x_1$, there is another ball centered at $x_1$ where $h = 0$. Therefore, we can 'slide' our center of the power series from $x_1$ to another $x_2$. If we can slide this to everywhere in $(a, b)$ we will get that the power series representation at every point is 0.

*optional material:* How do we guarantee that we can slide everywhere? This requires another idea from analysis called compactness. In short, the compactness of the interval $[x_0, \tilde{x}]$ or $[\tilde{x}, x_0]$ guarantees that we can slide our center of the power series across from $x_0$ to any $\tilde{x} \in (a, b)$ in a finite amount of steps.

We will give another way to prove this, as we have introduced Zorn's lemma. Without loss of generality, let $x_0 < \tilde{x} \in (a, b)$ Consider $S$, the set of points $x \leq \tilde{x}$ that you can 'slide to' from $x_0$ in a finite amount of steps. I claim that every increasing sequence in $S$ is bounded above by some element in $S$. Let $x_1 \leq x_2 \leq x_3 \leq \dots$ be an increasing sequence in $S$. We take $y = \lim_{n\to\infty} x_n$, then the series is bounded above by $y$. To construct the finite sequence going from $x_0$ to $y$, we see that $f$ is analytic at $y$ thus it has a power series representation centered at $y$ that converges to $f$ for some $B(y, \delta)$. We take $m$ large such that $x_m > y - \delta/4$. If we slide the power series centered at $y$ to be centered at $x_m$, the power series converges to $f$ at least in $B(x_m, 3\delta/4) \ni y$. That is, you can recenter the power series from $x_m$ to $y$. Therefore, we take the finite sequence that recenters the power series at $x_0$ to $x_m$, then recenter that sequence at $y$. Therefore $S$ contains a maximal element by Zorn's lemma. Finally, to find out what this maximal element is, we make use of the fact that the points where $f$ is analytic is open. Therefore, the only point that can be the maximal element of

$S$ is $\tilde{x}$, which is used as the upper limit of all elements in $S$. Therefore $\tilde{x}$ is the maximal element in $S$, thus $f = 0$ in some ball centered at $\tilde{x}$. We picked $\tilde{x}$ to be arbitrary, so $f = 0$ everywhere in $(a, b)$.

Corollary 1.9: Let $x_0 \in (a, b)$, and $f$ is analytic on $(a, b)$, $f$ equals the power series centered at $x_0$ where the power series converges.

*Proof.* The power series is analytic, and equals $f$ on some small open ball in $(a, b)$. □

## Exercises

1. Find the Taylor Series for the given functions at the indicated points.
   (a) $f(x) = e^{-x}, x_0 = 0$.
   (b) $f(x) = e^{x}, x_0 = 1$.
   (c) $f(x) = 1/x, x_0 = 1$.
   (d) $f(x) = \cos(x), x_0 = \pi/2$.
   (e) $f(x) = \ln(x), x_0 = 1$.
2. Determine the radius of convergence of the given function about $x = 0$.
   (a) $f(x) = (1 + x)/(x - 2)$.
   (b) $f(x) = 2x/(1 + 2x^2)$.
   (c) $f(x) = 1/(1 - t^3)$.
   (d) $f(x) = ((t - 4)(t^2 + 3))^{-1}$.

## 1.2 Taylor's Theorem with Remainder

Sometimes we don't want to take the whole power series representation, but truncate the series to get an approximation for the functions.

**TODO:** show that the n-th order taylor series is the best n-th order approximation for a function. As Taylor Series are used to approximate functions, it is of relevance to determine the accuracy of a series in representing its desired function.

**Definition 1.10 (Taylor's Formula with Remainder)**

The remainder of order n of the Taylor expansion of $f(x_o)$ is represented by the function,

$$R_n(x) = f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(x_o)}{k!}(x - x_o)^k.$$

The remainder of a Taylor expansion is the difference between the value of a function $f$ at $x$ and the partial sum of the $n^{th}$ term Taylor series. The series converges if $\lim_{n\to\infty} R_n = 0$.

**Theorem 1.11 ()**

Let $f(x)$ be a function on the interval $(a, b)$. $f$ is analytic on $(a, b)$ if there exists and $M > 0$ such that

$$|f^{(n)}(x) \leq M^n$$

for all $x \in (a, b)$ and $n \in \mathbb{N}$.

As a result of this theorem, the Taylor series expansion holds for all $x \in (a, b)$.

*Proof.* Let $f(x)$ be a function on the interval $(a, b)$ and $x_o \in (a, b)$. For some $M \in \mathbb{R}$ set $C = maxM|a - x_o, M|b - x_o|$. Then the $n^{th}$ term remainder of the Taylor expansion of $f(x)$ at $x_o$ is given by

$$R_n = f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(x_o)}{k!}(x - x_o)^k = \sum_{k=n}^{\infty} \frac{f^{(k)}(x_o)}{k!}(x - x_o)^k$$

Each term in this infinite series for $R_n$ is given by

$$R_k = \frac{f^{(k)}(x_o)}{k!}(x - x_o)^k \leq \frac{M^k}{k!}$$

## 1.3 The Binomial Theorem

The Binomial Theorem describes the expansion of powers of a binomial expression. A binomial is an algebraic expression consisting of two terms, such as $(a + b)$. The Binomial Theorem provides a way to expand expressions of the form $(a + b)^n$, where $n$ is a non-negative integer. The theorem can be stated as follows:

**Theorem 1.12 (Binomial Theorem)**

For any integer $n \geq 0$, the expansion of $(a + b)^n$ is given by:

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

where $\binom{n}{k}$ is the binomial coefficient, defined as:

$$\binom{n}{k} = \frac{n!}{k!(n - k)!}$$

for $k = 0, 1, 2, \ldots, n$.

In the expansion $(a + b)^n$, the sum consists of $n + 1$ terms, where each term has the form $\binom{n}{k} a^{n-k} b^k$. The key components of the expansion are:

- The *binomial coefficient* $\binom{n}{k}$, also called a combination, represents the number of ways to choose $k$ items from $n$ items, and is calculated using the formula:

$$\binom{n}{k} = \frac{n!}{k!(n - k)!}$$

- The powers of $a$ and $b$ decrease and increase respectively in each term, starting from $a^n$ for $k = 0$ to $b^n$ for $k = n$.
- The sum runs over all integer values of $k$ from 0 to $n$.

Example 1.13

Find the binomial expansion for $(a + b)^2$ where $a, b \in \mathbb{R}$.

Using the Binomial Theorem, we can expand $(a + b)^2$ as follows:

$$(a+b)^2 = \sum_{k=0}^{2} \binom{2}{k} a^{2-k} b^k$$

This gives the following terms:

$$= \binom{2}{0} a^2 b^0 + \binom{2}{1} a^1 b^1 + \binom{2}{2} a^0 b^2$$

Using the binomial coefficients:

$$= 1 \cdot a^2 + 2 \cdot ab + 1 \cdot b^2$$

Thus, the expanded form is:

$$(a+b)^2 = a^2 + 2ab + b^2$$

Example 1.14

Find the binomial expansion of $(x + y)^3$.

For $(x + y)^3$, we apply the Binomial Theorem:

$$(x+y)^3 = \sum_{k=0}^{3} \binom{3}{k} x^{3-k} y^k$$

This expands to:

$$= \binom{3}{0} x^3 y^0 + \binom{3}{1} x^2 y^1 + \binom{3}{2} x^1 y^2 + \binom{3}{3} x^0 y^3$$

Substituting the binomial coefficients:

$$= 1 \cdot x^3 + 3 \cdot x^2 y + 3 \cdot xy^2 + 1 \cdot y^3$$

Thus, the expanded form is:

$$(x+y)^3 = x^3 + 3x^2 y + 3xy^2 + y^3$$

Example 1.15

Find the binomial expansion of $(1 + \epsilon)^n$ for small $\epsilon$ and $n \in \mathbb{N}$.

For $(1 + \epsilon)^n$, we apply the Binomial Theorem with $a = 1$ and $b = \epsilon$:

$$(1+\epsilon)^n = \sum_{k=0}^{n} \binom{n}{k} 1^{n-k} \epsilon^k$$

This simplifies to:

$$= \binom{n}{0} + \binom{n}{1}\epsilon + \binom{n}{2}\epsilon^2 + \cdots + \binom{n}{n}\epsilon^n$$

The expansion provides the terms of $(1+\epsilon)^n$ for small values of $\epsilon$. Taking an approximation for small $\epsilon$, we can truncate the series to obtain an approximation.

$$(1+\epsilon)^n \approx 1 + n\epsilon.$$

This result is useful for many physical applications where small perturbations are considered.

### 1.3.1 Properties of Binomial Coefficients

The binomial coefficients $\binom{n}{k}$ have several important properties:

- $\binom{n}{0} = \binom{n}{n} = 1$
- $\binom{n}{k} = \binom{n}{n-k}$ (Symmetry property)
- The sum of the binomial coefficients for a given $n$ is:

  $$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

  This is known as the binomial identity.
- Pascal's identity:

  $$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

### 1.3.2 Negative and Non-Integer Exponents

The Binomial Theorem can also be extended to cases where the exponent $n$ is not a non-negative integer, although the series then becomes infinite. For example, if $n$ is a positive integer, the expansion of $(1+x)^n$ converges to a finite sum, but if $n$ is a negative integer or a fraction, the series may converge to an infinite sum. For $|x| < 1$, the Binomial Theorem for any real number $n$ can be written as:

$$(1+x)^n = \sum_{k=0}^{\infty} \binom{n}{k} x^k$$

where $\binom{n}{k}$ is generalized as:

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}$$

## 1.4 Multidimensional Taylor Series

The Taylor series is a powerful tool for approximating functions using polynomials. While the standard Taylor series applies to functions of a single variable, the *multidimensional Taylor series* extends this concept to functions of multiple variables. In this section, we will derive and explain the multidimensional Taylor series for a function of several variables. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a function that is sufficiently smooth (i.e., has continuous partial derivatives) in a neighborhood of a point

$\mathbf{a} = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$. The goal is to approximate $f$ near the point $\mathbf{a}$ using a polynomial expansion. Suppose we have a function $f(x_1, x_2, \ldots, x_n)$ of $n$ variables, the Multidimensional Taylor series expansion around a point $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ is given by

$$f(x_1, x_2, \ldots, x_n) = f(a_1, a_2, \ldots, a_n) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a_1, a_2, \ldots, a_n)(x_i - a_i) + \cdots$$

$$+\frac{1}{k!} \sum_{i_1, i_2, \ldots, i_k = 1}^{n} \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}}(a_1, a_2, \ldots, a_n)(x_{i_1} - a_{i_1})(x_{i_2} - a_{i_2}) \cdots (x_{i_k} - a_{i_k}) + \cdots$$

where the sums are taken over all possible combinations of partial derivatives of $f$.

**Definition 1.16 (Multidimensional Taylor Series)**

The general form of the Taylor series for a function $f\colon \mathbb{R}^n \to \mathbb{R}$ about a point $\mathbf{a}$ is

$$f(\mathbf{x}) = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{i_1, i_2, \ldots, i_k = 1}^{n} \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}}(\mathbf{a}) \prod_{j=1}^{k} (x_{i_j} - a_{i_j})$$

We will now examine the first and second-order approximations of the function using the Taylor series. The first-order approximation, also known as the linearization of the function $f$ around the point $\mathbf{a}$, is obtained by truncating the series after the first derivative term:

$$f(\mathbf{x}) \approx f(\mathbf{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})(x_i - a_i)$$

This approximation gives a linear model of the function near the point $\mathbf{a}$. The second-order approximation includes the terms up to the second derivative. It is given by:

$$f(\mathbf{x}) \approx f(\mathbf{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})(x_i - a_i) + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})(x_i - a_i)(x_j - a_j)$$

This provides a quadratic approximation to the function near $\mathbf{a}$, and is useful for understanding the curvature of the function around the point. Quadratic approximations are especially useful in physical applications as it indicates the stability of equilibria of dynamical systems.

**Example 1.17**

Compute the second-order Taylor approximation of $f(x, y) = e^{x^2+y^2}$ around the point $(0, 0)$.

First, compute the necessary partial derivatives:

$$f(x, y) = e^{x^2+y^2}$$

$$\frac{\partial f}{\partial x} = 2xe^{x^2+y^2}, \quad \frac{\partial f}{\partial y} = 2ye^{x^2+y^2}$$

$$\frac{\partial^2 f}{\partial x^2} = 2e^{x^2+y^2} + 4x^2 e^{x^2+y^2}, \quad \frac{\partial^2 f}{\partial y^2} = 2e^{x^2+y^2} + 4y^2 e^{x^2+y^2}$$

$$\frac{\partial^2 f}{\partial x \partial y} = 4xye^{x^2+y^2}$$

Evaluating these derivatives at $(0,0)$:

$$f(0,0) = e^0 = 1$$

$$\frac{\partial f}{\partial x}(0,0) = 0, \quad \frac{\partial f}{\partial y}(0,0) = 0$$

$$\frac{\partial^2 f}{\partial x^2}(0,0) = 2, \quad \frac{\partial^2 f}{\partial y^2}(0,0) = 2, \quad \frac{\partial^2 f}{\partial x \partial y}(0,0) = 0$$

Now, the second-order Taylor expansion around $(0,0)$ is:

$$f(x,y) \approx 1 + 0 \cdot (x-0) + 0 \cdot (y-0) + \frac{1}{2}\left[2(x-0)^2 + 2(y-0)^2\right]$$

$$f(x,y) \approx 1 + (x^2 + y^2)$$

Thus, the second-order approximation for $f(x,y) = e^{x^2+y^2}$ around $(0,0)$ is:

$$f(x,y) \approx 1 + x^2 + y^2$$

### 1.4.1 Error in the Taylor Expansion

The error in the Taylor series approximation is related to the remainder term $R_n(\mathbf{x})$, which represents the difference between the exact value of the function and its approximation up to the $n$-th degree. For a multivariable Taylor series, the remainder term can be written as:

$$R_n(\mathbf{x}) = \frac{1}{(n+1)!} \sum_{i_1, i_2, \ldots, i_{n+1}} \frac{\partial^{n+1} f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_{n+1}}}(\mathbf{a}) \prod_{j=1}^{n+1} (x_{i_j} - a_{i_j})$$

This error term quantifies the difference between the approximation and the exact value of the function.

## 1.5 Extrema of Multivariate Functions

Just as we are interested in finding the extreme points of functions of a single variable, we likewise wish to solve for stationary points of multivariate functions. Analysis of functions of multiple variables have analogous first and second derivative tests to those learned in single variable calculus. I used the term 'stationary point' as in three dimensions in addition to maxima or minima there exist so called saddle points. A 3D representation of a saddle point is presented in figure

**Definition 1.18 (Multivariable First Derivative Test)**

A point $(x_o, y_o) \in \mathbb{R}^{\not\vDash}$ is a *stationary point* of some function $f(x,y)$ if

$$\nabla f|_{(x,y)=(x_o,y_o)} = 0$$

Also similarly to the second derivative test in single variable calculus, we also have an analogous second derivative test in multivariable calculus to determine the classification of critical points. For this we use the discussion of multivariable Taylor series discussed in section . To second order, the Taylor expansion of some function $f(x, y)$ around $(x_o, y_o)$ is

$$f(x,y) \approx f(x_o,y_o) + \nabla f(x_o,y_o)^T d + \frac{1}{2!} d^T H f(x_o,y_o) d + R_2(x,y)$$

From the first derivative test $\nabla f(x_o, y_o) = 0$ thereby eliminating that term. Also, we know $R_2 \to 0, (x, y) \to (x_o, y_o)$. Thus we have

$$f(x,y) \approx f(x_o,y_o) + \frac{1}{2!} d^T H f(x_o,y_o) d$$

Rearranging we have

$$f(x,y) - f(x_o,y_o) = \frac{1}{2!} d^T H f(x_o,y_o) d$$

The left side appears as the numerator of the definition of the derivative. The sign of our derivative is dependent on the Hessian matrix $H$ which holds for all points $(x, y)$ near $(x_o, y_o)$.

## 1.6 Lagrange Multipliers

Lagrange multipliers are a method to find the extrema of a function subject to a constraint. Suppose we have a function $f(x, y)$ and a constraint $g(x, y) = c$. The method of Lagrange multipliers states that the extrema of $f(x, y)$ subject to the constraint $g(x, y) = c$ are found at points $(x, y)$ where the gradient of $f$ is parallel to the gradient of $g$. This implies

$$\nabla f(x,y) = \lambda \nabla g(x,y) \tag{1.1}$$

$$\begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \lambda \begin{bmatrix} \frac{\partial g}{\partial x} \\ \frac{\partial g}{\partial y}. \end{bmatrix} \tag{1.2}$$

The theory behind Lagrange multipliers is that we wish to find the points where $\nabla f \cdot g = 0$. This implies that the gradient of $f$ is at some maximum or minimum (or stationary point) when subject to $g$. We can use the system of equations (1.2) and our constraint equation to solve for our three unknowns: $\lambda$, $x$, and $y$. This result will gives us the extrema we are looking for. This method can be generalized to any $n$-dimensional function $f(\mathbf{x})$ subject to a constraint $g(\mathbf{x}) = c$. In this case we will have a system with $n + 1$ equations and $n + 1$ unknowns to solve for the extrema.

Example 1.19

Find the dimensions of the box with the largest volume such that its surface area is 100 square units.

Let $x$, $y$, and $z$ be the dimensions of the box. The volume of the box is $V(x, y, z) = xyz$ and the surface area is $S(x, y, z) = 2xy + 2yz + 2xz = 100$. We now apply our method of Lagrange Multipliers to maximize $V$ subject to $S$. We have the system of equations

$$\begin{bmatrix} yz \\ xz \\ xy \end{bmatrix} = \lambda \begin{bmatrix} 2y + 2z \\ 2x + 2z \\ 2x + 2y \end{bmatrix}$$
$$2xy + 2yz + 2xz = 100.$$

Solving this system of equations we find

$$x = y = z = 5$$

$$\lambda = \frac{1}{5}.$$

Therefore, the dimensions of the box with the largest volume are 5 units by 5 units by 5 units. The corresponding maximum volume is 125 cubic units.

# Chapter 2

# First Order Differential Equations

## 2.1 Introduction

A differential equation is an equation that relates an undetermined function with one or more of its derivatives. We call equations involving only single-variable derivatives of functions *ordinary differential equations* (ODEs) and those containing partial derivatives of multivariable functions *partial differential equations*(PDEs). We will focus on the former in this course and leave study of the latter to MATH 381.

The highest order derivative occurring in an ODE defines the *order* of the differential equation. We will look at first and second order ordinary differential equations. ODEs can be either homogeneous or inhomogeneous. Homogeneous equations have all terms involving the function $y$ or derivatives of $y$ summed to equal 0, while inhomogeneous equations will sum to equal a nonzero term.

$$\text{Homogeneous}: f(y, y', ..., y^n, t) = 0$$

$$\text{Inhomogeneous}: f(y, y', ..., y^n) = g(t)$$

Another classification of differential equations is concerned with the linearity of the terms. We can have either linear or nonlinear equations. An ODE

$$f(y, y', ...y^n, t) = g(t)$$

is linear if $f$ is linear with respect to terms involving the variable $y$ or derivatives of $y$. The general form looks like

$$a_0(t)y^n + .... + a_n(t)y = g(t)$$

Nonlinear equations will typically have terms involving $y$ or derivatives of $y$ multiplied together or terms involving nonlinear functions of $y$ such as $sin(y)$ or $e^y$.

## 2.2 Separation of Variables

A separable differential equation is any differential equation of the form,

$$N(y)\frac{dy}{dt} = M(t)$$

This allows us to multiply across by $dt$ and integrate both sides to find a function $y(t)$.

$$\int N(y(t))\frac{dy}{dt}dt = \int M(t)dt$$

I have written $N(y) = N(y(t))$ since $y$ is a function of $t$. Then we can suppose that $\frac{d}{dt}(y(t)) = N(y(t))\frac{dy}{dt}$. Which leads to the conclusion

$$y(t) = \int M(t)dt + C$$

for some constant $C$. You may have seen the differential treated as a fraction that can be separated and while that is sufficient for all computation purposes and will lead to the same answer, the formulation above is more mathematically rigorous.

## 2.3 Differential Forms

Let's take a look back at section (). The line integral of a vector field $F$ along a curve $C$ is defined as

$$\int_C F \cdot d\vec{r} = \int_C Mdx + Ndy$$

where $F = \langle M, N\rangle$. Given a the pair of functions $M(x, y)$ and $N(x, y)$, the expression

$$Mdx + Ndy$$

is called a *differential form*. Differential forms are generalizations of derivatives and integrals to manifolds. They are used to define integrals on curves, surfaces, volumes, and higher-dimensional geometries. A differenti 0-form is a smooth function. A differential 1-form is a linear map that takes a vector field as input and returns a function. In $\mathbf{R}^3$ this is of the form,

$$\omega = Pdx + Qdy + Rdz$$

where $P, Q, R$ are smooth functions of $x, y, z$. A differential 2-form is a linear map that takes two vector fields as input and returns a function. In $\mathbf{R}^3$ this is of the form,

$$\omega = Adxdy + Bdydz + Cdzdx.$$

In general, a differential $k$-form is a linear map that takes $k$ vector fields as input and returns a function. Expressed for $\mathbf{R}^n$ this is of the form,

$$\omega = \sum_{i_1 < i_2 < \ldots < i_k} A_{i_1, i_2, \ldots, i_k} dx_{i_1} dx_{i_2} \ldots dx_{i_k}.$$

In higher level mathematics, the wedge product $\wedge$ is used to define differential forms, however for our purposes how we have defined them is sufficient. Further study of differential forms is left to a differential geometry course.
We return to differential 2-forms in the context of first order differential equations. Suppose we have $F = \langle M, N\rangle$. Then,

$$Mdx + Ndy = F \cdot d\vec{r}.$$

We know that a vector field $F$ is conservative if it is the gradient of a scalar field $f$. That is, $F = \nabla f$. Then, $F \cdot d\vec{r} = df$. Thus, if $F$ is conservative, then $Mdx + Ndy = df$. This suggests,

$$\frac{\partial f}{\partial x} = M \quad \text{and} \quad \frac{\partial f}{\partial y} = N$$

such that,

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy.$$

Therefore, $F$ is conservative if and only if $Mdx + Ndy$ is equal to the differential of some function $f$. We define such forms as exact. Recall from section (), some vector field $F = \langle M, N \rangle$ is conservative if,

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}. \tag{2.1}$$

In terms of differential forms, this is equivalent to saying $Mdx + Ndy$ is closed. One must be careful here, every exact form is closed because all conservative vector fields satisfy 2.1, however not all closed forms are exact. That is, some vector fields may satisfy 2.1 but not be conservative.

## 2.4 Exact Equations

We now look to make use of the concept of differential forms to solve first order differential equations. Consider the differential equation,

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

We can rewrite this as,

$$Mdx + Ndy = 0$$

Suppose $F = \langle M, N \rangle$ is exact, then

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}.$$

This means there exists some function $f$ such that $\nabla f = F = \langle M, N \rangle$. Section () tells us that exact forms correspond to conservative vector fields. Recall from our study of line integrals that the line integral of a conservative vector field $F$ is $f(b) - f(a)$ where $\nabla f = F$ and $a$ and $b$ are the endpoints of the curve. Therefore, we can integrate our differential form

$$\int Mdx + Ndy$$

to find $f(x, y)$. In practice, we integrate $Mdx$ and $Ndy$ separately. This will gives us

$$\int Mdx + g(y)$$

and

$$\int Ndy + h(x).$$

where $g(y)$ and $h(x)$ are functions of $y$ and $x$ respectively. This is because the derivative of a function of $x$ with respect to $y$ is 0 and vice versa. We can put together the two resulting functions to find $f(x, y)$.

Example 2.1

$$\frac{dy}{dx} = -\frac{2x + y}{x + 2y}$$

We can rewrite this as,

$$(x + 2y)dy + (2x + y)dx = 0.$$

We will apply the screening test to see if this is an exact equation.

$$\frac{\partial M}{\partial y} = 1 \quad \text{and} \quad \frac{\partial N}{\partial x} = 1.$$

Therefore, this is an exact equation. We can integrate to find $f(x, y)$.

$$\int (x + 2y)dy = xy + y^2 + g(x)$$

and

$$\int (2x + y)dx = x^2 + xy + h(y).$$

We can put these together to find $f(x, y)$.

$$f(x, y) = xy + y^2 + x^2 + xy = x^2 + 2xy + y^2.$$

## 2.5 Integration Factors

The above formulation gives a useful technique for solving exact differential equations, however we must now consider the case where the equation is not exact. The technique we will use if called the method of integrating factors. The idea is to multiply the equation by some function $\mu(x, y)$ such that the resulting equation is exact. That is, we want to find $\mu(x, y)$ such that,

$$\mu(x, y)Mdx + \mu(x, y)Ndy = 0$$

with

$$\frac{\partial}{\partial y}(\mu M) = \frac{\partial}{\partial x}(\mu N).$$

Using chain rule we can rewrite this as,

$$\mu\frac{\partial M}{\partial y} + M\frac{\partial \mu}{\partial y} = \mu\frac{\partial N}{\partial x} + N\frac{\partial \mu}{\partial x}.$$

$$\mu[\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}] = N\frac{\partial \mu}{\partial x} - M\frac{\partial \mu}{\partial y}.$$

Unfortunately, this is a partial differential equation and is not easy to solve. However, we can make use of the fact that $\mu$ is a function of $x$ and $y$. We can rewrite the above equation as,

$$\frac{\partial \mu}{\partial x} = \frac{1}{N}[\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}]\mu$$
$$\frac{\partial \mu}{\partial y} = -\frac{1}{M}[\frac{\partial M}{\partial y} - \frac{\partial N}{\partial x}]\mu.$$

We can treat these as two separable first order differential equations.

$$\frac{\partial \mu}{\mu} = \int$$

## 2.6 Variation of Parameters

Following our discussion of first order homogeneous differential equations, we now move on to discussing methods of findings solutions to inhomogeneous first order differential equations.

$$\frac{dy}{dx} + a(x)y = b(x)$$

We propose a solution $y(x) = u(x)h(x)$ where $h(x)$ is the solution to the corresponding homogeneous equation.

$$\frac{dy}{dx} + a(x)y = 0$$

The solution to this equation is

$$h(x) = e^{-\int a(x)dx}.$$

Going back to our solution form $y(x) = u(x)h(x)$ and substituting into our inhomogeneous equation

$$\frac{du}{dx}h + u\frac{dh}{dx} + a(x)uh = b(x)$$
$$\frac{du}{dx}h + u\left(\frac{dh}{dx} + a(x)h\right) = b(x)$$

Since $h(x)$ is a solution to the homogeneous equation, the term in the parenthesis vanish. Therefore our differential equation becomes

$$\frac{du}{dx} = \frac{b}{h}$$

Solving for $u$ we get

$$u = \int \frac{b(x)}{h(x)}dx$$

Lastly, multiplying by $h(x)$ to get our full solution $y(x)$

$$y(x) = h(x)\left(\int \frac{b(x)}{h(x)}dx + C\right)$$

Notice here that I have already included the constant of integration here. This is because the method of solving inhomogeneous differential equations often settles down to combining a general and particular solution. We see that the constant multiplied by $h(x)$ will give us a general solution to the homogeneous equation while the product of the term in the integral and $h(x)$ will give a particular solution.

## 2.7 Existence and Uniqueness of Solutions for First-Order ODEs

The study of first-order ordinary differential equations often revolves around understanding whether a solution exists for a given equation and, if so, whether that solution is unique. This section presents the fundamental results regarding existence and uniqueness, along with illustrative examples. A first-order ODE is typically written as:

$$\frac{dy}{dx} = f(x, y), \quad y(x_0) = y_0, \tag{2.2}$$

where $f(x, y)$ is a given function, and $y(x_0) = y_0$ specifies an initial condition at $x = x_0$. The goal is to determine whether there exists a function $y(x)$ satisfying Equation (2.2) and whether this solution is unique in a neighborhood of $x_0$. The classical result addressing this problem is the Existence and Uniqueness Theorem, often attributed to Picard-Lindelöf. The theorem is stated as follows:

**Theorem 2.2 (Existence and Uniqueness)**

Let $f(x, y)$ be a continuous function defined on a rectangle $R = \{(x, y) : a \le x \le b, c \le y \le d\}$ in the $xy$-plane. Suppose further that $f(x, y)$ satisfies a Lipschitz condition in $y$; that is, there exists a constant $L > 0$ such that

$$|f(x, y_1) - f(x, y_2)| \le L|y_1 - y_2| \tag{2.3}$$

for all $(x, y_1), (x, y_2) \in R$. Then, for any point $(x_0, y_0) \in R$, there exists a unique solution $y(x)$ to the initial value problem (2.2) defined on some interval $x_0 - \delta < x < x_0 + \delta$, where $\delta > 0$.

*Proof.* Consider the function $y(x)$ defined and continuous on the interval $x_0 - \delta < x < x_0 + \delta$ for some $\delta > 0$ with the initial condition $y(x_0) = y_0$. Hence, the function $f(x, y(x))$ is well-defined and continuous on this interval. Suppose $f(x, y(x)) = y'(x)$. We can integrate both sides of the differential equation to obtain

$$y(x) = y(x_0) + \int_{x_0}^{x} f(t, y(x))dt. \tag{2.4}$$

Any solution $y(x)$ to 2.4 satisfies the original differential equation. We can differentiate both sides with respect to $x$ to recover the original differential equation. Now we move on to proving uniqueness. For any two numbers $a_1$ and $a_2$ in the interval $x_0 - \delta < x < x_0 + \delta$, it follows from the Mean Value Theorem that

$$\frac{f(s, a_1) - f(s, a_2)}{a_1 - a_2} = \frac{\partial f}{\partial y}(s, z), \tag{2.5}$$

for some $z$ between $a_1$ and $a_2$. Equation 2.5 can be rewritten as

$$|f(s, a_1) - f(s, a_2)| \le L|a_1 - a_2|, \tag{2.6}$$

for all $a_1, a_2$ and $x_1, x_2$ in the interval $x_0 - \delta < x < x_0 + \delta$.
In other words, the conditions of the theorem states,
1. The continuity of $f(x, y)$ ensures the existence of solutions. Intuitively, if $f$ is not continuous, the differential equation may exhibit abrupt changes that preclude the formation of a well-defined

solution.
2. The Lipschitz condition guarantees uniqueness. This condition implies that $f(x, y)$ does not change too rapidly with respect to $y$, preventing the trajectories of different solutions from crossing.

Example 2.3

Consider the initial value problem

$$\frac{dy}{dx} = 2x + 3y, \quad y(0) = 1.$$

Here, $f(x, y) = 2x + 3y$. Since $f(x, y)$ is a linear function of $y$, it satisfies the Lipschitz condition with $L = 3$. Furthermore, $f(x, y)$ is continuous everywhere. Hence, by the Existence and Uniqueness Theorem, there exists a unique solution.

Example 2.4

Consider the initial value problem

$$\frac{dy}{dx} = y^{1/3}, \quad y(0) = 0.$$

Here, $f(x, y) = y^{1/3}$ is continuous but does not satisfy the Lipschitz condition at $y = 0$. Multiple solutions exist, such as $y(x) = 0$ and $y(x) = \left(\frac{2}{3}x\right)^{3/2}$. This example highlights the necessity of the Lipschitz condition for uniqueness.

# Chapter 3

# Second Order Differential Equations

## 3.1 Introduction

Our study of differential equations continues with second order equations. Such equations are defined by their highest order derivative of the function in question being a second derivative. They are of the form

$$a(t)y'' + b(t)y' + c(t)y = f(t),$$

where $a(t)$, $b(t)$, $c(t)$, and $f(t)$ are functions of $t$. We will study methods to solve these equations, both homogeneous and inhomogeneous, and discuss the existence and uniqueness of solutions.

## 3.2 Constant Coefficients

The first technique we will study in solving second order differential equations is for cases of homogeneous equations with constant coefficients. Such equations are of the form

$$ay'' + by' + cy = 0$$

This equation suggests we are looking for solutions $y(t)$ for which the derivatives can be easily summed together to produce zero. Methods in calculus suggests the solution

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

Let's suppose this is the case, then

$$y'(t) = re^{rt}$$

$$y''(t) = r^2e^{rt}$$

Substituting these into our differential equation

$$a(r^2e^{rt}) + b(re^{rt}) + c(e^{rt}) = 0$$

$$(ar^2 + br + c)e^{rt} = 0$$

Since $e^{rt} \neq 0$, this implies we want to find values $r$ which satisfy

$$ar^2 + br + c = 0$$

The fundamental theorem of algebra states that solving this equation will produce at least one complex root and 2 roots total counted for multiplicity. In this section, we will look at this case in which the equation produces two roots distinct $r_1, r_2 \in \mathbb{R}$. Thus we get two solutions,

$$y_1 = e^{r_1 t}$$

$$y_2 = e^{r_2 t}$$

We check linear independence with the Wronskian,

$$\begin{vmatrix} e^{r_1 t} & e^{r_2 t} \\ r_1 e^{r_1 t} & r_2 e^{r_2 t} \end{vmatrix} = r_2 e^{r_1} e^{r_2} - r_1 e^{r_1 t} e^{r_2 t} = (r_2 - r_1) e^{(r_1 + r_2)t} \neq 0$$

since the exponential function is never zero and $r_1, r_2$ are distinct. This gives us two linearly independent solutions that produce a basis for the set of solutions to this differential equation, thus our general solution is

$$y = c_1 e^{r_1 t} + c_2 e^{r_2 t}$$

where $c_1, c_2$ are undetermined coefficients to be determined by initial conditions.

## Exercises

1. Solve the following homogeneous second-order differential equation with the given initial conditions:

$$y'' - 3y' + 2y = 0,$$
$$y(0) = 1, \quad y'(0) = 0.$$

2. Solve the following homogeneous second-order differential equation with the given initial conditions:

$$y'' + 4y' + 4y = 0,$$
$$y(0) = 0, \quad y'(0) = 1.$$

## 3.3 Complex Roots

We now look at cases of equations in the previous section for which the characteristic equation produces complex roots. However, a quick remark is needed first.

**Definition 3.1 (Make this a remark somehow)**

A polynomial of degree 2 with real coefficients can either have, 2 real, 2 complex, 1 repeated real or 1 repeated complex roots.

This means that any degree two polynomial cannot have one real root and one complex root. We will now look at cases for which we have two complex roots. Suppose we have a differential equation of the form

$$ay'' + by' + cy = 0$$

The previous section suggests we solve the quadratic equation

$$ar^2 + br + c = 0$$

to find $r_1$ and $r_2$ that produce solutions $y_1 = e^{r_1 t}$ and $y_2 = e^{r_2 t}$ to the differential equation. If

$$r_1 = a_1 + b_1 i$$

$$r_2 = a_2 + b_2 i$$

then our general solution becomes

$$y = c_1 e^{(a_1 + b_1 i)t} + c_2 e^{(a_2 + b_2 i)t}$$

However, certain cases prove it useful to find real solutions. In these cases we use Euler's Identity

$$e^{(a+bi)t} = e^{at}(cos(bt) + isin(bt))$$

. The power in this technique is that it produces two real solutions from a single complex solution. We will prove this now.

*Proof.* Suppose $y = u + iv$ is a complex solution to the second order homogeneous differential equation with constant coefficients

$$ay'' + by' + cy = 0$$

where $u$ and $v$ are real valued functions. We have

$$y' = u' + iv'$$

$$y'' = u'' + iv''$$

Substituting into our differential equation

$$a(u'' + iv'') + b(u' + iv') + c(u + iv) = 0$$

$$= (au'' + bu' + u) + i(av'' + bv' + cv) = 0$$

This suggests both the real and imaginary parts of this equation must be zero thus we have,

$$au'' + bu' + cu = 0$$

$$i(av'' + bv' + cv) = 0$$

This results suggests that both $u$ and $v$ are real solutions to the differential equation. Now if we let $u = cos(bt)$ and $v = sin(bt)$ we have obtained two real solutions to our differential equation from one complex solution. We still must have the $e^a t$ factor multiplied by $u + v$ and our two undetermined coefficients to be satisfied by initial conditions; therefore our general solution is

$$y = e^{at}(c_1 cos(bt) + c_2 sin(bt))$$

We check linear independence with the Wronskian

$$\left| \right|$$

From this result we see that for cases of complex roots only one root suffices to obtain a general solution.

## 3.4 Method of Reduction of Order

When our characteristic equations of second-order constant coefficient homogeneous equations results in repeated roots, we obtain only one solution. Therefore we look to develop a technique to find a second solution. We suggest a solution of the form

$$y(t) = v(t)y_1(t)$$

where $y_1(t)$ is the first solution found. Taking derivatives we have

$$y'(t) = v'(t)y_1(t) + v(t)y_1'(t)$$

$$y''(t) = v''(t)y_1(t) + 2v'(t)y_1'(t) + v(t)y_1''(t)$$

Substituting this into the differential equation to solve for the undetermined equation $v(t)$

$$ay'' + by' + cy = 0$$

$$a(v''y_1 + 2v'y_1' + vy_1'') + b(v'y_1 + v'y_1') + c(vy_1) = 0$$

$$av''y_1 + v'(2ay_1' + by_1) + v(ay_1'' + by_1' + cy_1) = 0$$

The third term is zero since $y_1$ is a solution to that differential equation which is the same as what we started out with. Therefore we have

$$av''y_1 + v'(2ay_1' + by_1) = 0$$

$$\frac{v''}{v'} = \frac{-(2ay_1' + by_1)}{ay_1}$$

We can solve this separable differential equation for $v'$

$$\int \frac{dv}{v} = \int (-\frac{2y_1'}{y_1} - \frac{b}{a})dt$$

This gives us the solution

$$v' = \frac{1}{y_1^2}Ce^{-\int \frac{b}{a}dt}$$

We have kept the argument in the exponential in integral form as this method is generalizable to any differential equation in which we have one solution and require another for a general solution, however, in the case of constant coefficients the exponential will be $e^{bt/a}$.

## Exercises

3. Given that $y_1(x) = e^x$ is a solution to the differential equation:

   $$y'' - y' = 0,$$

   use the method of reduction of order to find a second, linearly independent solution $y_2(x)$.

4. Suppose $y_1(x) = x$ is a solution to the differential equation:

   $$x^2y'' - 3xy' + 3y = 0 \quad \text{for } x > 0.$$

   Use the method of reduction of order to find a second solution $y_2(x)$ that is linearly independent of $y_1(x)$.

## 3.5 Variation of Parameters

Thus far we have developed techniques to solving homogeneous second order equations. We now turn our attention to finding methods to solve inhomogeneous equations.

$$ay'' + by' + cy = f(t)$$

We find solutions to inhomogeneous equations by adding a general solution to the corresponding inhomogeneous equation with a particular solution to the inhomogeneous equation.

$$y = y_g + y_p$$

Suppose we can find two linearly independent solutions to the corresponding homogeneous equation of (), $y_1, y_2$. The method of variation of parameters suggests we look for a particular solution of the form

$$y(t)_p = u(t)y_1(t) + u_2(t)y_2(t)$$

where $y_1(t), y_2(t)$ are solutions to the corresponding homogeneous equation and $u_1(t), u_2(t)$ are undetermined coefficients. Taking the derivative

$$y' = u_1'y_1 + u_1y_1' + u_2'y_2 + u_2y_2'$$

These calculations are made simpler if we set

$$u_1'y_1 + u_2'y_2 = 0$$

Therefore $y'$ becomes

$$y' = u_1y_1' + u_2y_2'$$

Finding $y''$

$$y'' = u_1'y_1' + u_1y_1'' + u_2'y_2' + u_2y_2''$$

Substituting this into our differential equation

$$a(u_1'y_1' + u_1y_1'' + u_2'y_2' + u_2y_2'') + b(u_1y_1' + u_2y_2') + c(u_1y_1 + u_2y_2) = f(t)$$

$$u_1(ay_1'' + by_1' + cy_1) + u_2(ay_2'' + by_2' + cy_2) + au_1'y_1' + au_2'y_2' = f(t)$$

Since $y_1$ and $y_2$ are solutions to the corresponding homogeneous equation, the first two terms are zero. Thus, our equation reduces to

$$au_1'y_1' + au_2'y_2' = f(t)$$

We now have two equations and two unknowns.

$$u_1'y_1 + u_2'y_2 = 0$$

$$au_1'y_1' + au_2'y_2' = f(t)$$

From this we obtain

$$u_1' = \frac{-y_2f(t)/a}{y_1y_2' - y_1'y_2}$$

$$u_2' = \frac{y_1f(t)/a}{y_1y_2' - y_1'y_2}$$

We can integrate to find $u_1$ and $u_2$.

$$u_1 = \int \frac{-y_2 f(t)/a}{y_1 y_2' - y_1' y_2} dt$$

$$u_2 = \int \frac{y_1 f(t)/a}{y_1 y_2' - y_1' y_2} dt$$

You may notice that the argument in the denominator is the Wronksian thereby implying that if our solutions $y_1, y_2$ are not linearly independent, then when don't have the requisite information to form a general solution to the differential equation. It that case we must return to section 3.5 and find a second linearly independent equation via Method of Reduction of Order.

This provides us with our particular solution to the inhomogeneous differential equation

$$y_p = y_1 \int \frac{-y_2 f(t)/a}{y_1 y_2' - y_1' y_2} dt + y_2 \int \frac{y_1 f(t)/a}{y_1 y_2' - y_1' y_2} dx$$

We add this to the general solution to the corresponding homogeneous equation to obtain the full solution to the inhomogeneous equation.

$$y(x) = y_g + y_p$$

Example 3.2

Find the general solution to the differential equation

$$y'' - 3y' + 2y = 2e^x$$

The characteristic equation of the corresponding homogeneous equation is

$$r^2 - 3r + 2 = 0$$

This equation has roots $r_1 = 1$ and $r_2 = 2$. Thus the general solution to the corresponding homogeneous equation is

$$y_g = c_1 e^x + c_2 e^{2x}$$

We can use equation () compute the particular solution to the differential equation.

$$y_1 = e^x$$

$$y_2 = e^{2x}$$

$$y_1' = e^x$$

$$y_2' = 2e^{2x}$$

$$\begin{aligned} y_p &= e^x \int \frac{-e^{2x} 2e^x}{e^x 2e^{2x} - e^{2x} e^x} dx + e^{2x} \int \frac{e^x 2e^x}{e^x 2e^{2x} - e^{2x} e^x} dx \\ &= e^x \int \frac{-2e^{3x}}{e^x 2e^{2x} - 2e^{3x}} dx + e^{2x} \int \frac{2e^{2x}}{e^x 2e^{2x} - 2e^{3x}} dx \\ &= e^x \int \frac{-2e^{3x}}{2e^{3x}} dx + e^{2x} \int \frac{2e^{2x}}{2e^{3x}} dx \end{aligned}$$

$$= e^x \int -1dx + e^{2x} \int e^{-x}dx$$
$$= -e^x + e^{2x}(-e^{-x})$$
$$= -e^x - e^x$$
$$= -2e^x$$

Thus our particular solution is

$$y_p = -2e^x$$

Therefore our full solution to the differential equation is

$$y(x) = c_1e^x + c_2e^{2x} - 2e^x.$$

## Exercises

5. Solve the non-homogeneous differential equation:

$$y'' + y = \sin(x),$$

using the method of variation of parameters. Assume that the complementary solution is given by:

$$y_c(x) = C_1 \cos(x) + C_2 \sin(x).$$

6. Use the method of variation of parameters to solve the non-homogeneous differential equation:

$$x^2y'' - 3xy' + 3y = x^3,$$

for $x > 0$. Assume that the complementary solution is:

$$y_c(x) = C_1x + C_2x^3.$$

## 3.6 Method of Undetermined Coefficients

There are certain classes of inhomogeneous equations such that we can propose a solution form and algebraically solve for specifying parameters. Such classes usually involve equations of constant coefficients and inhomogeneous terms of familiar functions like exponentials or sinusoidals. As in the previous section, to find a full solution to an inhomogeneous equation we sum together a general solution to the corresponding homogeneous equation with a particular solution to the inhomogeneous equation. Suppose we have the second order inhomogeneous differential equation

$$a(t)y'' + b(t)y' + c(t)y = Acos(\omega t) + Bsin(\omega t)$$

We propose a particular solution of the form

$$y_p = X_1cos(\omega t) + X_2sin(\omega t)$$

Taking derivatives

$$y' = -\omega X_1sin(\omega t) + \omega X_2cos(\omega t)$$
$$y'' = -\omega^2X_1cos(\omega t) - \omega^2X_2sin(\omega t)$$

Substituting into our differential equation

$$a(-\omega^2 X_1 cos(\omega t) - \omega^2 X_2 sin(\omega t)) + b(-\omega X_1 sin(\omega t) + \omega X_2 cos(\omega t) + c(X_1 cos(\omega t) + X_2 sin(\omega t))$$

$$= Acos(\omega t) + Bsin(\omega t)$$

We rearrange to get a single cosine and sine term on each side

$$(-a\omega^2 X_1 + b\omega X_2 + cX_1)cos(\omega t) + (-a\omega^2 X_2 - b\omega X_1 + cX_2)sin(\omega t)$$

$$= Acos(\omega t) + Bsin(\omega t)$$

From this it is apparent that

$$(-a\omega^2 + c)X_1 + b\omega X_2 = A$$

$$(-a\omega^2 + c)X_2 - b\omega X_1 = B$$

We can solve this system of equations to find

$$X_1 = \frac{aB - bA}{a^2\omega^2 + b^2\omega^2 - c}$$

$$X_2 = \frac{aA + bB}{a^2\omega^2 + b^2\omega^2 - c}$$

Thus our particular solution is

$$y_p = \frac{aB - bA}{a^2\omega^2 + b^2\omega^2 - c} cos(\omega t) + \frac{aA + bB}{a^2\omega^2 + b^2\omega^2 - c} sin(\omega t)$$

We add this to the general solution to the corresponding homogeneous equation to obtain the full solution to the inhomogeneous equation.

$$y(x) = y_g + y_p$$

where $y(g)$ will dependent on the form of our differential equation. Now suppose we have an inhomogeneous equation of the form

$$a(t)y'' + b(t)y' + c(t)y = e^{\alpha t}$$

We propose a particular solution of the form

$$y_p = Xe^{\alpha t}$$

Taking derivatives

$$y' = \alpha X e^{\alpha t}$$

$$y'' = \alpha^2 X e^{\alpha t}$$

Substituting into our differential equation

$$a\alpha^2 X e^{\alpha t} + b\alpha X e^{\alpha t} + cXe^{\alpha t} = e^{\alpha t}$$

$$X(a\alpha^2 + b\alpha + c) = 1$$

$$X = \frac{1}{a\alpha^2 + b\alpha + c}$$

Thus our particular solution is

$$y_p = \frac{1}{a\alpha^2 + b\alpha + c} e^{\alpha t}$$

We add this to the general solution to the corresponding homogeneous equation to obtain the full solution to the inhomogeneous equation.

$$y(x) = y_g + y_p$$

where $y(g)$ will once again be dependent on the form of our differential equation.

Example 3.3

Find the general solution to the differential equation

$$y'' + 4y = 2cos(2t)$$

Since we have a sinusoidal inhomogeneous term, we propose a particular solution of the form

$$y_p = X_1 cos(2t) + X_2 sin(2t)$$

Taking derivatives

$$y' = -2X_1 sin(2t) + 2X_2 cos(2t)$$

$$y'' = -4X_1 cos(2t) - 4X_2 sin(2t)$$

Substituting into our differential equation

$$-4X_1 cos(2t) - 4X_2 sin(2t) + 4X_1 cos(2t) + 4X_2 sin(2t) = 2cos(2t)$$

$$= 2cos(2t)$$

This implies that $X_1 = 1/2$ and $X_2 = 0$. Therefore our particular solution is

$$y_p = \frac{1}{2} cos(2t)$$

The general solution to the corresponding homogeneous equation is

$$y_g = c_1 cos(2t) + c_2 sin(2t)$$

Thus our general solution to the differential equation is

$$y(x) = c_1 cos(2t) + c_2 sin(2t) + \frac{1}{2} cos(2t)$$

Example 3.4

Find the general solution to the differential equation

$$y'' + 4y = 2e^{2t}$$

Since we have an exponential inhomogeneous term, we propose a particular solution of the form

$$y_p = Xe^{2t}$$

Taking derivatives

$$y' = 2Xe^{2t}$$
$$y'' = 4Xe^{2t}$$

Substituting into our differential equation

$$4Xe^{2t} + 4Xe^{2t} = 2e^{2t}$$
$$= 2e^{2t}$$

This implies that $X = 1/4$. Therefore our particular solution is

$$y_p = \frac{1}{4}e^{2t}$$

The general solution to the corresponding homogeneous equation is

$$y_g = c_1 cos(2t) + c_2 sin(2t)$$

Thus our general solution to the differential equation is

$$y(x) = c_1 cos(2t) + c_2 sin(2t) + \frac{1}{4}e^{2t}$$

## Exercises

7. Solve the following non-homogeneous differential equation using the method of undetermined coefficients:
$$y'' - 3y' + 2y = e^{2x},$$
subject to the initial conditions $y(0) = 0$ and $y'(0) = 1$.

8. Solve the following non-homogeneous differential equation using the method of undetermined coefficients:
$$y'' + y = \sin(x),$$
subject to the initial conditions $y(0) = 1$ and $y'(0) = 0$.

## 3.7 Existence and Uniqueness

# Chapter 4

# Eigenvalues and Eigenvectors

## 4.1 Definition of Eigenvectors and Eigenvalues

We look to examine the behavior of linear transformations in which a vector space maps to itself. We denote $T \in \mathcal{L}(V)$ as the linear transformation $T : V \to V$ where $\mathcal{L}(V)$ is the set of all operators $\mathcal{L}(V, V)$. In order to perform operations on a subspace $U$ of $V$, we look to define a special class of operators that maps $U$ to itself.

**Definition 4.5 (Invariant Subspaces)**

Suppose $U$ is a subspace of $V$. $U$ is *invariant* for a given transformation $T : V \to V$, if $Tu \in U$ for any $u \in U$.

Vectors that constitute invariant subspaces and their change under $T$ are specially defined.

**Definition 4.6 (Eigenvalues and Eigenvectors)**

Suppose $U \in V$ is invariant under $T$ and $u$ is a nonzero vector in $U$. Then,

$$Tu = \lambda u$$

where $\lambda \in \mathbb{F}$ is the *eigenvalue* of $T$ and $u$ is it's corresponding *eigenvector*.

It is important to note that for a given eigenvalue there may be multiple eigenvectors. The dimension of the subspace the eigenvectors for a given eigenvalue span (called the *eigenspace*) corresponds to the number of eigenvectors for the given eigenvalue.
Rewriting the () gives,

$$(T - \lambda I)u = 0.$$

By construction it is apparent that the set of eigenvectors of $T$ is equal to $null(T - \lambda I)$. Since we have a nonzero vector mapping to zero, one can see that $\lambda$ is an eigenvalue of $T$ if and only if $T - \lambda I$ is not injective. And, since this gives a noninvertible square matrix by SOME THEOREM $\lambda$ is an eigenvalue of $T$ if and only if $T - \lambda I$ is not surjective as well.
By SOME THEOREM, the determinant of a noninvertible matrix is zero. This property allows us to solve for the value of $\lambda$.

Theorem 4.7 ()

Suppose $\lambda_1, \lambda_2, ..., \lambda_m$ are distinct eigenvalues of $T : V \to V$ corresponding to distinct eigenvectors $u_1, u_2, ..., u_m$. Then the eigenvectors $u_1, u_2, ..., u_k$ are linearly independent.

*Proof.* We proceed by contradiction. Suppose $u_1, u_2, ..., u_m$ are linearly dependent. Choose $k$ to be the smallest integer such that,

$$u_k \in span\{u_1, u_2, ..., u_{k-1}\}.$$

Therefore $u_k$ can be written as,

$$u_k = a_1u_1 + a_2u_2 + ... + a_{k-1}u_{k-1}.$$

Take the transformation $T$ of both sides of the equation,

$$Tu_k = T(a_1u_1 + a_2u_2 + ... + a_{k-1}u_{k-1})$$

$$\lambda_k u_k = a_1\lambda_1 u_1 + a_2\lambda_2 u_2 + ... + a_{k-1}\lambda_{k-1}u_{k-1}.$$

Multiply both sides of () by $\lambda_k$ and subtract () to obtain,

$$0 = a_1(\lambda_k - \lambda_1)u_1 + ... + a_{k-1}(\lambda_k - \lambda k - 1)u_{k-1}.$$

By construction, this implies that $a_i = 0$ for $i \in (1, k-1)$ since the eigenvectors are linearly independent and the eigenvalues are distinct. However, this implies $u_k = 0$, a contradiction since we don't consider $\vec{0}$ and eigenvector.

Corollary 4.8: There are at most $n$ distinct eigenvalues for each operator on an $n$-dimensional vector space.

Therefore, suppose we have $T \in \mathcal{L}(V)$ with $n$ distinct eigenvalues, then it follows that $T$ has $n$ distinct eigenvectors. From the previous theorem the set of eigenvectors to $T$ must be linearly independent therefore $n \leq dim(V)$.

## 4.2 Computing Eigenvalues and Eigenvectors

We look to develop a method to solve for the eigenvalues and eigenvectors of some transformation $T \in \mathcal{L}(V, V)$. Suppose $T(x) = Ax$ and $n = dim(V)$, this implies that $A$ is $n \times n$. We look for $\lambda \in \mathbf{F}$ that satisfies

$$Ax = \lambda x.$$

Right multiplying each side by the identity matrix $n \times n$ identity matrix $I_n$ gives

$$Ax = \lambda Ix.$$

Solving to isolate $x$ produces the homogeneous equation

$$(A - \lambda I)x = 0.$$

From the previous section we know the eigenvectors of $A$ span $null(A)$. Therefore, we look for non-trivial vectors $x$ that solve $(A - \lambda I)$. This implies that $(A - \lambda I)$ must be non-invertible. We use the property that for non-invertible matrices the determinant is zero to solve for $\lambda$.

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

Computing the determinant of $(A - \lambda I)$ produces a polynomial $P_k(\lambda)$ where $k \le n$.

$$P_k(\lambda) = 0$$

Solving for the roots of $P_k(\lambda)$ finds the desired eigenvalues for $A$. For a polynomial of degree $k \le n$, there will be at most $k$ eigenvalues. We substitute each computed eigenvalue into $(A - \lambda I)x = 0$ to solve for vectors $x$ that span $null(A)$. Each $x$ is an eigenvector of $A$. The space spanned by each eigenvalue $\lambda$ is called the *eigenspace* of $\lambda$.

Example 4.9

Find the eigenvalues and eigenvectors of
$A = \begin{bmatrix} 1 & 4 & 3 \\ 4 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$.

We wish to find $\lambda$ that satisfy,

$$\begin{bmatrix} 1 & 4 & 3 \\ 4 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} x = \lambda x$$

. Algebraically rearranging,

$$(\begin{bmatrix} 1 & 4 & 3 \\ 4 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} - \lambda I)x = 0$$

$$\begin{bmatrix} 1-\lambda & 4 & 3 \\ 4 & 1-\lambda & 0 \\ 3 & 0 & 1-\lambda \end{bmatrix} x = 0$$

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

$$\begin{vmatrix} 1-\lambda & 4 & 3 \\ 4 & 1-\lambda & 0 \\ 3 & 0 & 1-\lambda \end{vmatrix} = (1-\lambda)((1-\lambda)^2 - 0) - 4(4(1-\lambda) - 0) + 3(0 - 3(1-\lambda))$$

$$= (1-\lambda)^3 - 25(1-\lambda) = (1-\lambda)((1-\lambda)^2 - 25)$$
$$= (1-\lambda)(\lambda^2 - 2\lambda - 24) = (1-\lambda)(6-\lambda)(4+\lambda) = 0$$

Therefore our eigenvalues are $\lambda = 1, 6$ and $-4$. We substitute each eigenvalue into $(A - \lambda I)x = 0$ to find the eigenvectors of $A$. For $\lambda = 1$,

$$(A - 1(I))x = \begin{bmatrix} 0 & 4 & 3 \\ 4 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} x = 0$$

By Gaussian-Jordan Reduction we get,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3/4 \\ 0 & 0 & 0 \end{bmatrix}$$

This means our eigenvector $\vec{x}$ is,

$$\vec{x} = x_3 \begin{bmatrix} 0 \\ -3/4 \\ 1 \end{bmatrix}$$

Repeating the same process for $\lambda = 6$,

$$(A - 6I) = \begin{bmatrix} -5 & 4 & 3 \\ 4 & -5 & 0 \\ 3 & 0 & -5 \end{bmatrix}.$$

By Gauss-Jordan Reduction we get,

$$\begin{bmatrix} 1 & 0 & -5/3 \\ 0 & 1 & -4/3 \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore our eigenvector is,

$$\vec{x} = x_3 \begin{bmatrix} 5/3 \\ 4/3 \\ 1 \end{bmatrix}$$

Lastly for $\lambda = 4$,

$$(A + 4I) = \begin{bmatrix} 5 & 4 & 3 \\ 4 & 5 & 0 \\ 3 & 0 & 5 \end{bmatrix}$$

. Gauss-Jordan Reduction gives us,

$$\begin{bmatrix} 1 & 0 & 5/3 \\ 0 & 1 & -4/3 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus our eigenvector is,

$$\vec{x} = x_3 \begin{bmatrix} -5/3 \\ 4/3 \\ 1 \end{bmatrix}.$$

So our set of eigenvectors for $A$ is,

$$\left\{ \begin{bmatrix} 0 \\ -3/4 \\ 1 \end{bmatrix}, \begin{bmatrix} 5/3 \\ 4/3 \\ 1 \end{bmatrix}, \begin{bmatrix} -5/3 \\ 4/3 \\ 1 \end{bmatrix} \right\}.$$

**Theorem 4.10 ()**

Suppose $A$ is an upper triangular matrix. Then the eigenvalues of $A$ are the the entries along the diagonal. Similarly, if $A$ were lower triangular the same result holds.

This theorem follows from the determinant of a triangular matrix being the product of the diagonal entries. Therefore, if we can subtract some $\lambda$ such that one of the entries becomes zero, then the matrix determinant is zero and the value of that $\lambda$ satisfies $Ax = \lambda x$.

# Exercises

## 4.3 Diagonalization

A diagonal matrix is a matrix consisting of only diagonal entries. The diagonal matrix for an operator $T \in \mathcal{L}(V)$ is

$$\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

if,

$$\begin{aligned} Tv_1 &= \lambda_1 v_1 \\ Tv_2 &= \lambda_2 v_2 \\ &\vdots \\ Tv_n &= \lambda_n v_n. \end{aligned}$$

This suggests $\lambda_i$ are the eigenvalues of $T$ and $v_i$ are the corresponding eigenvectors. What this tells us is that an arbitary operator has some diagonal matrix consisting of eigenvalues with respect to some basis of eigenvectors. Our goal is to find the basis of eigenvectors for some operator $T$ such that $T$ is diagonal.

**Theorem 4.11 (The Diagonalization Theorem)**

An operator $T \in \mathcal{L}(V)$ is diagonalizable if there exists a basis of $V$ consisting of eigenvectors of $T$. In matrix form, suppose $T(x) = Ax$ for some $n \times n$ matrix $A$. Then $A$ is diagonalizable if there exists an invertible matrix $P$ with the eigenvectors of $A$ as columns such that

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

is a diagonal matrix.

This theorem tells us that a $n \times n$ matrix $A$ is diagonalizable if the eigenvectors of $A$ form a basis for $\mathbf{R}^n$.

*Proof.* Let $A$ be a $n \times n$ matrix. Suppose $P$ is a $n \times n$ matrix with column vectors $v_i$ and $D$ be a $n \times n$ diagonal matrix with diagonal entries $\lambda_i$. Then,

$$AP = A \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$$

and,

$$PD = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 & \lambda_2 v_2 & \cdots & \lambda_n v_n \end{bmatrix}.$$

Suppose $\lambda_i$ are the eigenvalues of $A$ and $v_i$ are the corresponding eigenvectors. Then we obtain the following result,

$$AP = PD.$$

Since $P$ is invertible, we can multiply each side by $P^{-1}$ to obtain,

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

Therefore $A$ is diagonalizable. We have shown that a matrix $A$ is diagonalizable if there exists an invertible matrix $P$ such that $D = P^{-1}AP$ is a diagonal matrix. We will now develop the methodological steps to diagonalize a matrix $A$.

1. Find the eigenvalues of $A$ by solving $det(A - \lambda I) = 0$.
2. Find the corresponding eigenvectors of $A$ by solving $(A - \lambda I)v = 0$.
3. Form the matrix $P$ with the eigenvectors of $A$ as columns.
4. Form the matrix $D$ with the eigenvalues of $A$ as diagonal entries.

This will give us the diagonalization of $A$ as $A = PDP^{-1}$.

Example 4.12

Diagonalize the matrix $A = \begin{bmatrix} 1 & 4 & 3 \\ 4 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$.

We have already found the eigenvalues and eigenvectors of $A$ in the previous section. The eigenvalues of $A$ are $\lambda = 1, 6, -4$ with corresponding eigenvectors

$$\begin{bmatrix} 0 \\ -3/4 \\ 1 \end{bmatrix}, \begin{bmatrix} 5/3 \\ 4/3 \\ 1 \end{bmatrix}, \begin{bmatrix} -5/3 \\ 4/3 \\ 1 \end{bmatrix}.$$

Therefore the matrix $P$ is,

$$P = \begin{bmatrix} 0 & 5/3 & -5/3 \\ -3/4 & 4/3 & 4/3 \\ 1 & 1 & 1 \end{bmatrix}.$$

The matrix $D$ is,

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -4 \end{bmatrix}.$$

Therefore the diagonalization of $A$ is,

$$\begin{bmatrix} 1 & 4 & 3 \\ 4 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 5/3 & -5/3 \\ -3/4 & 4/3 & 4/3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} \frac{4}{125} & -\frac{3}{100} & \frac{3}{100} \\ \frac{3}{100} & \frac{2}{25} & \frac{2}{25} \\ -\frac{3}{100} & \frac{3}{100} & \frac{3}{100} \end{bmatrix}$$

Here I have skipped the calculation of $P^{-1}$ as it is quite tedious and one can use methods from section () to compute it.

## 4.4 Spectral Theorem

**Definition 4.13 (Self Adjoint Operators)**

An operator $T \in \mathcal{L}(V)$ is *self-adjoint* if $T = T^*$.

**Definition 4.14 (Hermitian Matrices)**

A matrix $A$ is *Hermitian* if $A = A^*$. If $A$ is real, then $A = A^T$. In this case, $A$ is symmetric.

## 4.5 Generalized Eigenvectors

From section () we know that an $n \times n$ matrix $A$ with $n$ distinct eigenvalues $\lambda_i$ has $n$ corresponding eigenvectors $\vec{v}_i$ which form a basis for $\mathbf{R}^n$ . In this case each $\lambda_i$ has algebraic and geometric multiplicities both equal to 1. However consider the matrix

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

It has one eigenvalue $\lambda = 1$ which has one corresponding eigenvector

$$\vec{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

We can see the eigenvectors of $A$ do not form a basis for $\mathbf{R}^2$. The algebraic multiplicity of $\lambda = 1$ is 2, however its geometric multiplicity is only 1. Therefore we see that matrices with a set of eigenvectors which do not form a basis for the column space of $A$ display an inequality between the geometric and algebraic multiplicities. We can generalize the above example to the following definition.

**Definition 4.15 (Defective Matrices)**

A $n \times n$ matrix is defective if the sum of its eigenvalues' algebraic multiplicities $\mu_a$ and the sum of geometric multiplicities $\mu_g$ has the property

$$\mu_a > \mu_g$$

The eigenvectors of defective matrices do not form a linearly independent basis for $\mathbf{R}^n$. This implies that such matrices are *non-diagonalizable*. However in cases for which we wish to diagonalize a matrix, compute a matrix exponential or find a basis consisting of eigenvectors, we seek a method to resolve this issue.

**Definition 4.16 (Generalized Eigenvectors)**

For a matrix $A$, some $\lambda \in \mathbf{F}$ is a *level $j$ generalized eigenvector* if it satisfies

$$(A - \lambda I)^j x = 0$$

Theorem 4.17 ()

Suppose $\lambda$ is an eigenvalue of $A$ with multiplicity $m$. Then

$$(A - \lambda I)^j = 0$$

with $j > m$ has the same solution space as

$$(A - \lambda I)^m = 0$$

## 4.6 Matrix Exponentials

Suppose we have the linear system of differential equations

$$\frac{dx}{dt} = Ax.$$

Our study of differential equations suggests a solution of the form

$$x = Ce^{At}.$$

Now we are presented with the problem of computing the exponential of a matrix. Let's take the series expansion of $e^x$.

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + ... + \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Choosing instead to expand the matrix $A$ we get,

$$e^A = 1 + A + \frac{A^2}{2!} + \frac{A^3}{3!} + ... + \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

We know how to compute powers of matrices however doing it in practice and taking high enough powers in the series to get a sufficient solution is quite futile (and an approximation on top of that). There are two methods building on concepts introduced earlier this chapter which allows us to exactly compute the matrix exponential, diagonalization and generalized eigenvectors. We will begin with the former which is the preferrable route.

Suppose there exists a matrix $P$ and a diagonal matrix $D$ such that,

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

for some matrix $A$. We can substitute this into () to obtain,

$$e^A = 1 + (PDP^{-1}) + \frac{(PDP^{-1})^2}{2!} + \frac{(PDP^{-1})^3}{3!} + ... + \sum_{n=0}^{\infty} \frac{(PDP^{-1})^n}{n!}.$$

From section 4.3, we know,

$$A^n = PD^nP^{-1}$$

. Therefore, our series expansion of $e^A$ becomes,

$$e^A = 1 + PDP^{-1} + \frac{PD^2P^{-1}}{2!} + \frac{PD^3P^{-1}}{3!} + ... + \sum_{n=0}^{\infty} \frac{PD^nP^{-1}}{n!}.$$

Factoring out $P$ and $P^{-1}$ we obtain,

$$e^A = P(1 + D + \frac{D^2}{2!} + \frac{D^3}{3!} + ... + \sum_{n=0}^{\infty} \frac{D^n}{n!})P^{-1}.$$

The expansion in the parenthesis is the matrix exponential of the diagonal matrix $D$, therefore,

$$e^A = Pe^D P^{-1}.$$

Since the exponential of a diagonal matrix is the matrix consisting of the diagonal entries exponetiated, equation () allows us to easily compute $e^A$.

However, in cases for which $A$ is non-diagonalizable we must find an alternative method to solve $e^A$. We will use generalized eigenvectors to do this.
Generalized Eigenvectors tell us that for some $\lambda_j$, the matrix $(A - \lambda_j I)^j$ is zero. This implies that for all $m > j$, $(A - \lambda_j I)^m = 0$. Therefore, suppose $\lambda$ is a jth level eigenvector. If we write $e^{(A-\lambda I)t}$ as a series expansion,

$$e^{(A-\lambda I)v} = 1 + (A - \lambda I)v + \frac{(A - \lambda I)^2 v^2}{2!} + \frac{(A - \lambda I)^3 v^3}{3!} + ... + \sum_{n=0}^{j-1} \frac{(A - \lambda I)^n v^n}{n!},$$

this series terminates at $A^{j-1}$. This is exactly what we are looking for in order to compute $e^A$.

## 4.7 The Fundamental Solution of a Matrix

Suppose we have a system of coupled differential equations described by:

$$\begin{aligned} x_1' &= a_{11}x_1 + \cdots + a_{1n}x_n \\ x_2' &= a_{21}x_1 + \cdots + a_{2n}x_n \\ &\vdots \\ x_n' &= a_{n1}x_1 + \cdots + a_{nn}x_n \end{aligned}$$

Where $x_1, ..., x_n$ are functions of $t$ with derivatives $x_1', ..., x_n'$ and $a_{ij}$ are constants.

$$\begin{aligned} A &= \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \\ \vec{x}(t) &= \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} \\ \vec{x'}(t) &= \begin{bmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{bmatrix} \\ A\vec{x}(t) &= \vec{x'}(t) \end{aligned}$$

We solve this equation by finding an $\vec{x}(t)$ that satisfies it on some interval of $t$.

**Definition 4.18 (Fundamental Solution of a Matrix)**

For an $n \times n$ matrix $A$, there exists a set of $n$ linearly independent functions $\vec{x_1}(t), ..., \vec{x_n}(t)$ which constitute an $n$-dimensional basis for the vector space of all solutions of $A$. We call this set of functions the **fundamental solution of the matrix** $A$.

Suppose we have the system of uncoupled differential equations

$$x_1'(t) = ax_1(t)$$

$$x_2'(t) = bx_2(t)$$

for some constants $a$ and $b$. This can be written in matrix form as

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix}$$

These equations suggest the solutions to this system of differential equations are

$$x_1(t) = c_1 e^{at}$$

$$x_2(t) = c_2 e^{bt}$$

From this we suggest the solution to any linear system of differential equations $A\vec{x} = \vec{x'}$ is of the form

$$\vec{x}(t) = \vec{v} e^{\lambda t}.$$

Taking the derivative $\vec{x'}(t)$

$$\vec{x'}(t) = \lambda \vec{v} e^{\lambda t}$$

Taking equation () once more and multiplying each side by $A$

$$A\vec{x}(t) = A\vec{v} e^{\lambda t}$$

The left sides of equations () and () are our differential equation thus our right sides must equal.

$$A\vec{v} e^{\lambda t} = \lambda \vec{v} e^{\lambda t}$$

This suggests that vectors $v$ and scalars $\lambda$ which satisfy this system of differential equations are eigenvectors and eigenvalues of the matrix $A$.

**Example 4.19**

$$\frac{dx}{dt} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} x$$

This matrix gives the eigenvalues $\lambda = 3, -1$ corresponding to the eigenvectors

$$\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \}$$

Therefore our solutions to the systems of differential equations are

$$\{c_1\vec{x_1}(t) = e^{3t}\begin{bmatrix}1\\1\end{bmatrix}, c_2\vec{v_2}(t) = e^{-t}\begin{bmatrix}-1\\1\end{bmatrix}\}$$

If we have vector solutions $\vec{x_1}(t), \vec{x_2}(t)\dots\vec{x_n}(t)$ to the system of differential equations $A\vec{x} = \vec{x'}$ then the matrix $\vec{X}(t)$ with columns $\vec{x_1}(t), \vec{x_2}(t)\dots\vec{x_n}(t)$ is the fundamental solution of the matrix $A$. This matrix $\vec{X}(t)$ is the linear combination of vector solutions to the system of differential equations.

$$\vec{X}(t) = \vec{x_1}(t) + \vec{x_2}(t) + \dots + \vec{x_n}(t)$$
$$\vec{X}(t) = \begin{bmatrix}\vec{x_1}(t) & \vec{x_2}(t) & \dots & \vec{x_n}(t)\end{bmatrix}$$

Therefore our fundamental solution $\vec{X}(t)$ to example 4.14 is,

$$\vec{X}(t) = \begin{bmatrix}e^{3t}\begin{bmatrix}1\\1\end{bmatrix} & e^{-t}\begin{bmatrix}-1\\1\end{bmatrix}\end{bmatrix}.$$