Contents

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.

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

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

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

□

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 $x_1 \leq x_2 \leq x_3 \leq \ldots$ 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 Multidimesnional Taylor Series

We look to extend the formulation for the Taylor series to approximate functions of multiple variables.

Definition 1.12 ()

1.4 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 1.13 (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).

1.5 Lagrange Multipliers

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

Differential forms are generalized methods for describing derivatives and integrals in multi-dimensional spaces. Suppose we have the differential equation

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

We can convert this equality into the form

$$Mdx + Ndy = 0$$

setting $f(x, y) = \frac{-M(x,y)}{N(x,y)}$. Solutions to each of these forms while conveying the same information, can be interpreted in different ways. A solution $f(x)$ to () is a collection of functions satisfying the differential equation, each corresponding to a different initial condition. A solution $f(x, y) = c$ to () is a collection of level curves in $\mathbf{R}^2$. Each $f(x)$ is a subset of one of the level curves $f(x, y) = c$.

Along a curve C, we can write $d\vec{r} = \langle dx, dy\rangle$ which points in the tangential direction. Similarly, we can define $F = \langle M, N\rangle$. Thus equation () can be rephrased as $F \cdot d\vec{r}$.

2.4 Integration Factors

2.5 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.

2.6 Existence and Uniqueness

Chapter 3

Second Order Differential Equations

3.1 Introduction

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,

$$y_1 = e^{r_1t}$$

$$y_2 = e^{r_2t}$$

We check linear independence with the Wronskian,

$$\begin{vmatrix} e^{r_1t} & e^{r_2t} \\ r_1e^{r_1t} & r_2e^{r_2t} \end{vmatrix} = r_2e^{r_1}e^{r_2} - r_1e^{r_1t}e^{r_2t} = (r_2 - r_1)e^{r_1+r_2)t} \neq 0$$

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.

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

$$| \; |$$

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

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

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 $e^{bt/a}$.

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

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

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

We can integrate to find u_1 and u_2.

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

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

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

$$y_p = y_1 \int \frac{-y_2 f(t)/a}{y_1y_2' - y_1'y_2}dt + y_2 \int \frac{y_1 f(t)/a}{y_1y_2' - y_1'y_2}dt$$

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

$$y' = -\omega X_1sin(\omega t) + \omega X_2cos(\omega t)$$

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

Substituting into our differential equation

$$a(-\omega^2 X_1cos(\omega t) - \omega^2 X_2sin(\omega t)) + b(-\omega X_1sin(\omega t) + \omega X_2cos(\omega t) + c(X_1cos(\omega t) + X_2sin(\omega t))$$

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

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 using linear algebra

$$\begin{bmatrix} -a\omega^2 + c & b\omega \\ -b\omega & -a\omega^2 + c \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} A \\ B \end{bmatrix}$$

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 4.2 (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 4.3 (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 λ 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 λ 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 λ.

Theorem 4.4 ()

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,

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

$$\lambda_k u_k = a_1\lambda_1u_1 + a_2\lambda_2u_2 + ... + a_{k-1}\lambda_{k-1}u_{k-1}.$$

Multiply both sides of () by λ_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 4.5: 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 λ.

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

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

$$P_k(\lambda) = 0$$

Solving for the roots of $P_k(\lambda)$ finds the desired eigenvalues for A. For a polynomial of degree $k \leq n$, 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 λ is called the *eigenspace* of λ.

Example 4.6

Consider $A = \begin{bmatrix} 1 & 4 & 3 \\ 4 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$. We wish to find λ 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,

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

Example 4.7

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 4.8 ()

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 λ such that one of the entries becomes zero, then the matrix determinant is zero and the value of that λ satisfies $Ax = \lambda x$.

Exercises

4.3 Matrix Exponentials

The

4.4 The Eigenvalue Method to Solving Ordinary Differential Equations

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

$$x_1' = a_{11}x_1 + ...a_{1n}x_n$$

$$x_2' = a_{21}x_1 + ...a_{2n}x_n$$

$$\vdots$$

$$x_n' = a_{n1}x_1 + ...a_{nn}x_n$$

Where $x_1, ..., x_n$ are functions of t with derivatives $x_1', ..., x_n'$ and a_{ij} are constants. We can write this system in terms of matrix vectors.

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

$$\vec{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$$

$$\vec{x'}(t) = \begin{bmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}$$

$$A\vec{x}(t) = \vec{x'}(t)$$

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

Definition 4.9 (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

$$x_1'(t) = 3x_1(t)$$

$$x_2'(t) = 2x_2(t)$$

This can be written in matrix form as

$$\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix}$$

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

$$x_1(t) = c_1 e^{3t}$$

$$x_2(t) = c_2 e^{2t}$$

From this we deduce that 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 λ which satisfy this system of differential equations are eigenvectors and eigenvalues of the matrix A.

Example 4.10

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

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

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

4.5 Generalized Eigenvectors

From section () we know that an $n \times n$ matrix A with n distinct eigenvalues λ_i has n corresponding eigenvectors $\vec{v_i}$ which form a basis for $\mathbf{R}^n$. In this case each λ_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 λ 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 4.11 (Non-Diagonalizable Matrix)