Numerical Methods for Solving Differential Equations

Ryan Hausner

July 2024

Contents

1 Background and Theory **2**

2 Numerical Methods for ODEs **4**
2.1 Euler's Method 5
2.2 Local Truncation Error 7
2.3 Global Truncation Error 9

3 An Improvement to Euler's Method **10**
3.1 Heun's Method 11

Abstract

Differential equations are a fundamental pillar in the study of mathematics. They appear everywhere: in the temperature of a cup of tea, the spread of disease, and population growth. They are a powerful tool to make models when predicting phenomena. Unfortunately, not all differential equations have simple solutions and several methods have been developed to numerical solve said equations. Some of these include Euler's Method, Heun's Method, and the Runge-Kutta Methods. The main objectives of this paper are to outline the different numerical methods for solving differential equations and compare them using MATLAB. Each method will be applied to multiple differential equations and the error between the numerical solution and the analytic solution will be calculated and charted.

Introduction

Differential equations play a vital role in engineering and science, modeling various systems. They provide a foundation for mathematically determining the behavior of said systems. They allow for the prediction of future behaviors based on current data.

Despite their importance in the real world, they don't always have simple solutions. An analytical solution to a differential equation involves using mathematical formulas which involve the differential equation to possess specific properties, often not found in the real world.

To combat this inadequacy, numerical solutions to differential equations have been developed. Numerical solutions provide approximations for the solutions to differential equations with no analytical solution or an analytic solution that is very hard to compute.

This paper aims to analyze the different numerical methods for approximating the solutions to Ordinary Differential Equations (ODEs) such as Euler's Method, Heun's Method, and the Runge-Kutta Methods for ODEs.

Rigorous comparative analysis will be applied to explore the strengths and weaknesses of each method and figure out where each method is best fit in real world applications.

1 Background and Theory

An ordinary differential equation is an equation that involves a function and its derivative. Differential equations describe how a function changes with respect to a variable (usually time). An example of an ODE is

$$\frac{dy}{dt} = f(t, y) \tag{1}$$

As seen in Equation 1, the function $f(t, y)$ is related to its derivative dy/dt.

When solving for differential equations, there are two different types of solutions. The first is the general solution. This is considered a "family" of solutions since there are many functions that satisfy the differential equation. These consist of constants that represent the family of functions that satisfy the equation. The second kind of solution is the particular solution. This is found when given an initial condition such as $y(0) = 2$. This can then be used with the general solution to solve for the value of the constants in the general solution.

To compare the general and particular solutions to differential equations, consider the following:

$$\frac{dy}{dt} = -ky^1 \tag{2}$$

To solve this differential equation, start by dividing by sides by t and integrate both sides with respect to t.

$$\frac{1}{y}\frac{dy}{dt} = -k$$

$$\int \frac{1}{y}\frac{dy}{dt}dt^2 = \int -kdt$$

$$ln(y) = -kt + c$$

$$y = C_0 e^{-kt} \tag{3}$$

Equation 3 is the general solution to the differential equation $\frac{dy}{dt} = -ky$ (Equation 2). It represents the family of functions that are the solution to Eq. 2.

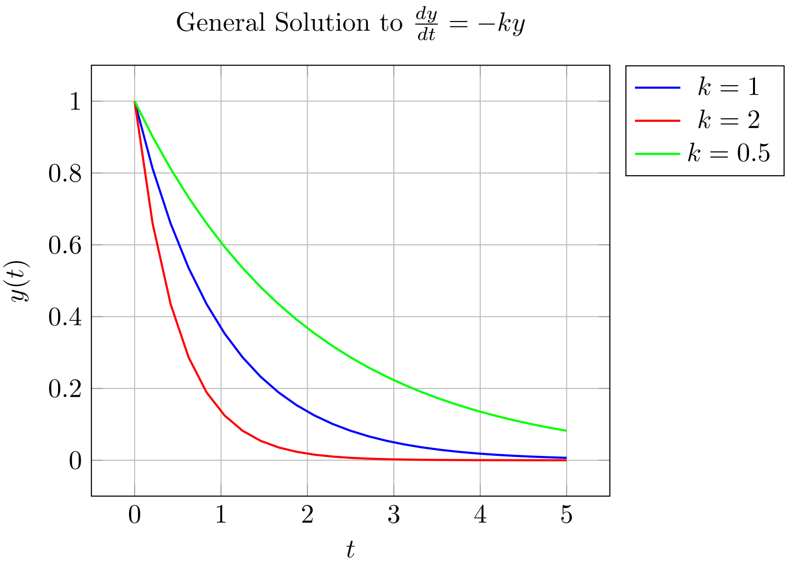

Figure 1: Graphs of the general solution $y(t) = C_0 e^{-kt}$ for different values of k.

To find the particular solution to Eq. 2, an initial condition needs to be present. If the initial conditions $y(0) = 2$ and $y(2) = 5$ are given, a particular solution can be found.

Plugging in the first condition$y(0) = 2$ the value for C_0 can be found to be 2.

[1]This is one of the most important differential equations and has many applications that will be considered later in the paper.

[2]The dt does not cancel. Instead, dy represents $\frac{dy}{dt}dt$.

$$y(t) = 2e^{-kt} \tag{4}$$

Plugging in the second initial condition $y(2) = 5$ the value for k can be found to be -0.4581.

Now that values for both constants C_0 and K have been found, the particular solution for Eq. 2 given the initial conditions $y(0) = 2$ and $y(2) = 5$ can be written.

$$y(t) = 2e^{-0.4581t} \tag{5}$$

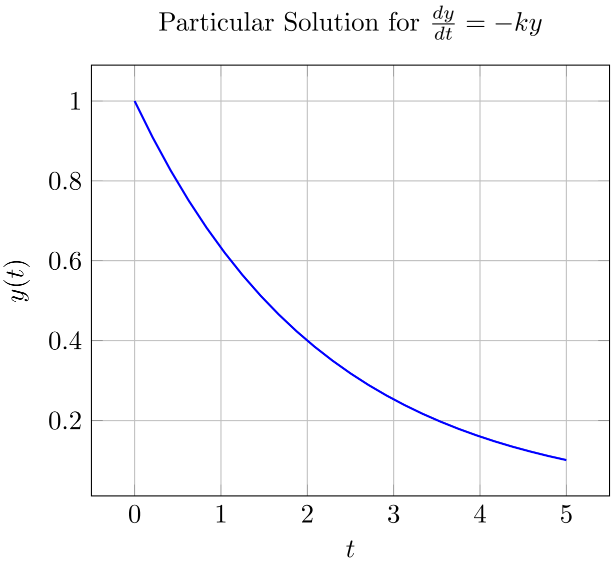

Figure 2: Graph of the particular solution $y(t) = C_0e^{-kt}$ given initial conditions $y(0) = 2$ and $y(2) = 5$.

This differential equation may represent many phenomena. One example is the growth rate of a species over time. The species may start with 2 members ($y(0) = 2$) and may end up with 5 members after 2 days ($y(2) = 5$). This demonstrates how powerful differential equations are as a tool for modeling many real world systems.

2 Numerical Methods for ODEs

Several methods for numerically solving differential equations have been developed. The three methods that will be further studied in this paper are Euler's Method, Heun's Method, and the Runge-Kutta Methods. Before comparing the three methods, it is important to understand the derivation and applications of each.

2.1 Euler's Method

Euler's Method is a simple, yet powerful tool for solving differential equations. To derive Euler's Method, recall the limit definition of the derivative.

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{6}$$

Instead of taking the limit as h approaches 0, Euler's Method takes a very small h value and uses that to approximate the derivative. Since the limit definition of the derivative includes both the derivative and the function itself, it allows computation of numerical solutions to differential equations that would be difficult to solve otherwise.

$$\frac{dy}{dt} \approx \frac{y(t+h)-y(t)}{h}$$

Rearranging the equation allows to find a new numerical value based on the previous value. The more iterations of Euler's Method, the more accurate the approximation becomes. The formula for Euler's Method is as follows.[3]

$$y(t_n) + \left.\frac{dy}{dt}\right|_{t=t_n} * h = y(t_{n+1}) \tag{7}$$

In words, this formula reads that the next "y" value is equal to the previous known "y" value plus some change (the derivative at the known point times a small change h). This is very similar to a tangent line approximation of a value from single variable calculus. [4]

Consider the differential equation $\frac{dy}{dt} = -0.5y$ with the initial condition $y(0) = 1$. This example can be solved analytically using the method previously discussed but rather than doing that we will use Euler's Method for a numerical approximation of the solution of $y(1)$. In this case the step-size (h) will be 0.5 (a fairly large step-size).

$$y(0.5) \approx y(0) + \left.\tfrac{dy}{dt}\right|_{y=0} * (0.5)$$

Following the first iteration, the new value ($y(0.5)$) becomes 0.875.

$$y(1) \approx y(0.5) + \left.\tfrac{dy}{dt}\right|_{y=0.5} * (0.5)$$

Doing two iterations of Euler's Method approximates the solution $y(1)$ to the differential equation $\frac{dy}{dt} = -0.5y$ as 0.5625.

Numerical methods of solving differential equations become a more accurate approximation the more iterations done. In this case, two iterations is very few and gives only a very rough approximation.

[3] Can also be written as $f(x_{n+1}) = f(x_n) + f'(x) * h$

[4] $y_{n+1} = y_n + \frac{dy}{dx}(\Delta x)$

Iterations	Step-Size	Approximation	Error $(y(1)_{approx} - y(1)_{actual})$
2	0.5	0.5625	-0.04404
3	0.33	0.5837	-0.02283
4	0.25	0.5861	-0.02043
5	0.2	0.5905	-0.01603
10	0.1	0.5987	-0.00783
100	0.01	0.6058	-0.00076
1000	0.001	0.6065	-0.00003

Table 1: Euler's Method with Different Iterations to approximate $y(1)$ for the differential equation $\frac{dy}{dt} = -0.5y$

As seen in Table 1, the more iterations done, the more accurate the approximation gets. [5]

Since this is a differential equation that can be solved analytically, it is important to compare the exact answer with the approximation found using Euler's Method. The general solution for this differential equation is $y(t) = e^{-0.5t}$. Plugging in for $y(1)$ yields $e^{-0.5}$ or 0.60653. As seen by the exact solution to this differential equation for $y(1)$, the numerical approximation found was very accurate with more iterations. Even with such a simply differential equation such as the one used in the previous example, Euler's Method proves to be very powerful in calculating numerical solutions.

While the last example was very simple to calculate analytically, not every differential equation is. Consider the following equation:

$$\frac{dy}{dt} = y^2 - t^2. \tag{8}$$

This differential equation is not as intuitive as the last, and will involve more manipulation in order to obtain an analytic solution. However, Euler's Method is easily applicable here. Consider Eq. 8 with the initial condition $y(0) = 1$. Using a large step size such as 1/2 ($h = 0.5$), a very rough approximation for numerical solutions would be easy to compute.

Iteration	t	y	$\frac{dy}{dt}$
1	0	1	1
2	0.5	1.5	2
3	1	2.5	5.25
4	1.5	5.125	24.0156
5	2	17.1328	289.533

Table 2: Euler's Method calculations for $\frac{dy}{dt} = y^2 - t^2$ with step size $h = 0.5$.

[5] The error between the numerical approximation and the analytical solution will be discussed in a later section.

With a few simple calculations, we were able to obtain an approximation to solutions to a differential equation that would otherwise be difficult to solve. Because of the small step size, there is a greater margin of error between the approximation and the exact solution. For differential equation where the derivative changes at a more volatile rate, a smaller step size is preferred to diminish the error. With a large step size, the approximation becomes flawed by using the previous tangent line when the derivative changes by a great amount.

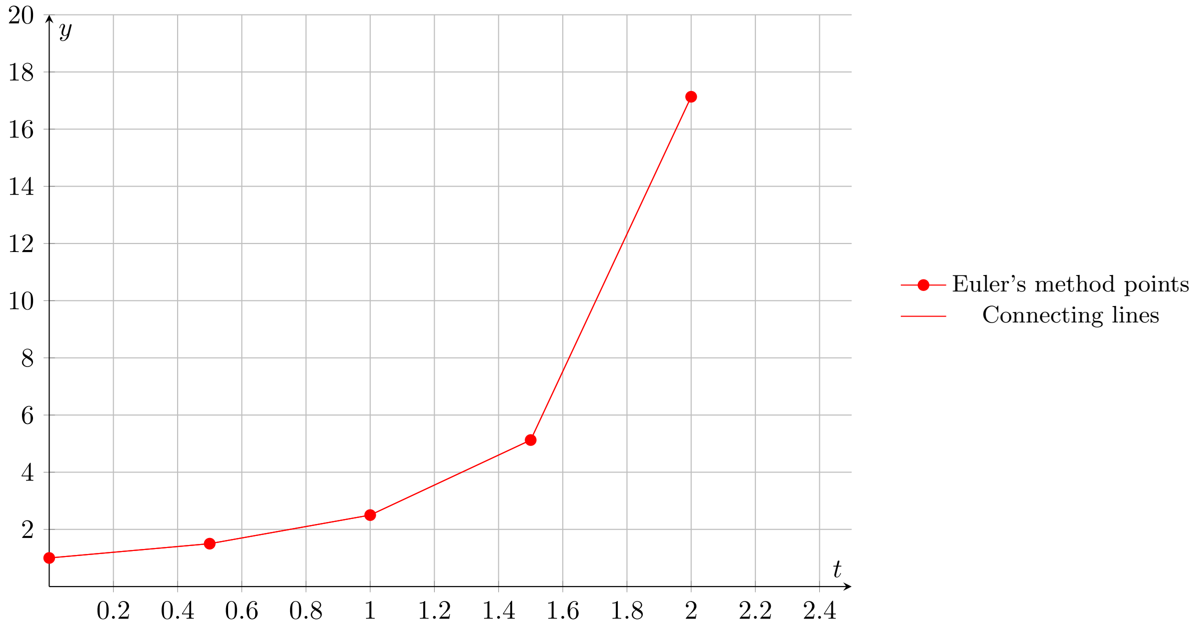

Figure 3: Euler's method for solving $\frac{dy}{dt} = y^2 - t^2$ with $y(0) = 1$ and step size $h = 0.5$. The plot shows the approximate solution using Euler's method.

As seen in Fig. 3, Euler's Method creates a model of the solution to the differential equation. If we were to decrease the step size h, this approximation would more closely fit the solution to the differential equation. [6] When h decreases, the tangent lines become infinitesimally small and form to the curvature of the exact solution.

2.2 Local Truncation Error

While Euler's Method is a very powerful tool to quickly compute approximations of solutions to differential equations, there is always an error between the actual answer and the approximation Euler's Method yields. The two types of error when dealing with Euler's Method are the Local Truncation Error (LTE) and the Global Truncation Error (GTE). To more clearly understand these errors, we will begin by demonstrating LTE.

When working with Euler's Method, it is first assumed that no previous steps have been done, i.e. you are starting with the initial condition given. The first type of error that is involved in Euler's

[6] Assuming many significant figures are used during calculation.

Method is the Local Truncation Error or the error produced by one iteration. This error is equal to the difference between the approximation and the true solution. Consider the differential equation and the following conditions:

$$\omega'(t) = f(t, \omega(t)) \tag{9}$$

$$\omega(t_n) = y_n \tag{10}$$

$$y_{n+1} = y_n + h * f(t_n, y_n) \tag{11}$$

For these conditions, $\omega(t_n)$ is the exact solution and Euler's Method must be applicable. If these conditions are met, the Local Truncation Error is the difference between the exact solution and the approximation found by one iteration of Euler's Method. The Local Truncation Error is given by:

$$LTE = \omega(t_{n+1}) - y_{n+1} \tag{12}$$

This error can be further explored by considering the 3rd degree Taylor Series for $\omega(t_{n+1})$ centered about t_n.

$$\omega(t_{n+1}) = \omega(t_n) + h\omega'(t_n) + \frac{h^2}{2}\omega''(t_n) + \frac{h^3}{3!}\omega'''(t_n)$$

Substituting this in Eq. 12 for $\omega(t_{n+1})$ and $\omega(t_n) + h\omega'(t_n)$ for y_{n+1} gives us:

$$[\omega(t_n) + h\omega'(t_n) + \frac{h^2}{2}\omega''(t_n) + \frac{h^3}{3!}\omega'''(t_n)] - [\omega(t_n) + h\omega'(t_n)]$$

$$LTE = \frac{h^2}{2}\omega''(t_n) + \frac{h^3}{3!}\omega'''(t_n)$$

By using the Taylor Expansion we obtain a more elegant way of expressing the error between the exact answer and the approximation.[7] This new way of expressing the Local Truncation Error shows that the error is h^2 times some unknown constant. The way we write this is:

$$LTE = Nh^2 + Mh^3 \tag{13}$$

Where N and M are unknown constants. From Eq. 13, the importance of a small h value becomes more apparent. If an h value of 0.5 was used compared to an h value of 1, the error will be approximately one fourth with half of the h value.

To summarize, the Local Truncation Error(LTE), or the error from a single step in Euler's Method is the difference between the exact solution and the approximation found. The error is inversely proportional to the h value squared times some constant.

[7] Although the Taylor Series continues, the terms after the 3rd degree are negligible.

2.3 Global Truncation Error

While the Local Truncation Error provides the error done by a single step h, the Global Truncation Error provides the error done throughout the process of Euler's Method.

Contrary to the formulaic way of defining the error done by a single step from the last section, it is simpler to describe the Global Error numerically.

If the error done by a single step is defined by:

$$\epsilon_{n+1} = \omega(t_n) - y_n \tag{14}$$

The error over the global domain propagates from the previous steps done. The error from the next step is dependent on the error from the previous step. The Global Truncation Error is not simply the sum of the Local Truncation Error but rather a reflection of how the error increases based on the error from the previous steps.

Consider the differential equation

$$\tfrac{dy}{dt} = -2y, y(0) = 1, t \in [0, 1]$$

Using Euler's Method with $h = 0.2$, find $y(1)$.

From our previous knowledge of differential equations, we know that the exact solution to this differential equation is $y(t) = e^{-2t}$.

$$y(0.2) \approx y(0) + (0.2)(-2) = 0.6$$

$$...$$

$$y(1) \approx y(0.8) + (0.2)(-0.2592) = 0.07776$$

If we first find the Local Truncation Error for the first iteration of the method, we find that $\omega(t_n) - y_n = 0.6703 - 0.6 = 0.0703$. This is an example of the Local Truncation Error. Now that we are doing more than one step, the Global Truncation Error can describe the error found as the error propagation from previous steps. Using the exact solution for $y(1)$, we find that $\epsilon_{t_1} = 0.0578$. While this may initially seem counterintuitive, it tells us about the behavior of the differential equation. Since the GTE at a later step is lower than the LTE, it shows that the approximation is becoming closer to the exact solution. This makes sense in the context of the problem we are working in. The equation $\frac{dy}{dt} = -2y$ has the exact solution $y(t) = e^{-2t}$ given the initial condition $y(0) = 1$. $y(t)$ is a function that rapidly decreases which lines up with the lower GTE in step 5 compared to the LTE in the first step. By only using numeric calculations, it is clear that the error propagation is less effective with a smaller h value.[8]

[8]There is a general formula for the Global Truncation Error that is derived and explained in Appendix A.

Now that Euler's Method has been explained and demonstrated along with Local Truncation Error and Global Truncation Error, it calls into question its efficacy and time to be used. It is evident that in order for an accurate approximation, there needs to be a small h. However, a small h value means that more iterations need to be done, which means more calculations have to be done. Euler's Method involves a balance between finding an h value that provides a fairly accurate approximation but is not too computationally heavy. This idea will be further explored in comparison to the other numerical methods for solving differential equations.

3 An Improvement to Euler's Method

The second method of numeric calculation is Heun's Method. This is sometimes called the improved Euler's Method because of the smaller error it produces. While Euler's Method is a very powerful and simple tool to approximate solutions to differential equations, it has its setbacks. Firstly, because it depends on the derivative from the previous point, it can't account for differential equations that change more sharply. When an equation is volatile and doesn't behave in a way that is simple, the GTE shows.

Consider the previous example:

$$\frac{dy}{dt} = y^2 - t^2, y(0) = 1$$

Recall that this has no simple analytical solution and therefore needs numerical methods. However, Euler's Method fails to account for the volatile nature of this equation and therefore leads to a very large error, ϵ especially with larger step-sizes. The previous section established that the smaller the h value, the more accurate the approximation.

H Value	$y(1)$ Approx	Error ϵ
0.5	2.5	-24.1888
0.2	3.6732	-23.0156
0.1	5.1644	-21.5244
0.01	15.345	-11.3438
0.001	24.53	-2.1588
0.0001	26.448	-0.2408

Table 3: Euler's Method calculations for $\frac{dy}{dt} = y^2 - t^2$ with varying step size.

While the other examples were easily approximated using Euler's Method, this one shows its inaccuracy until it reaches a very small step-size. This continues the issue of finding an h value that yields an accurate approximation without being computationally intensive.

3.1 Heun's Method

The next method to be explored is Heun's Method. Heun's Method is sometimes referred to as the revised or "improved" version to Euler's Method. It takes the foundation of Euler's Method and corrects it using a correction step. While this sounds more rigorous, it produces a more accurate approximation while using less steps.

The first step in Heun's Method is same as Euler's Method.

$$y_{n+1} = y_n + h * f(t_n, y_n)$$

Heun's Method then takes this new point using the estimation found by Euler's Method, (t_{n+1}, y_{n+1}) and uses that to calculate the slope at the new point.

$$y_{n+1} = y_n + \frac{h}{2}[f(t_n, y_n) + f(t_{n+1}, y_{n+1})] \tag{15}$$

This is Heun's Method. It takes the average of the two slopes in order to create a more accurate approximation. By using Euler's Method as an intermediate step, it is able to approximate solutions using less iterations.

Using the previous example from Euler's Method:

$$\frac{dy}{dt} = y^2 - t^2, y(0) = 1$$

H Value	$y(1)$ Euler Approximation	$y(1)$ Heun Approximation	ϵ Euler	ϵ Heun
0.5	2.5	4.6936	-24.1888	-21.9952
0.2	3.6732	8.8682	-23.0156	-17.8206
0.1	5.1644	13.7378	-21.5244	-12.951
0.01	15.345	25.9203	-11.3438	-0.7658
0.001	24.53	26.6779	-2.1588	-0.0109
0.0001	26.448	26.687	-0.2408	-0.0018

Table 4: Comparison of Euler's Method and Heun's Method for $\frac{dy}{dt} = y^2 - t^2$ with varying step size.