Imperial College London

Mathematics: Year 2

Numerical Analysis

Xin Wang

October 1, 2020

Abstract

Equations that represent mathematical models in engineering involve derivatives (**differential equations**) and integrals (**integral equations** or **integro-differential equations**) of the variables associated with the models.

While some differential equations can be solved, the vast majority cannot be solved, only approximations can be made. There are many techniques for analyzing the various differential equations. Many differential equations cannot be solved exactly thus will require numerical solutions to approximate the equations. Each numerical solution technique has varying levels of error in approximation under different circumstances so choosing the right one is important.

Contents

Listings

1 Arithmetic and error analysis

1.1 Exponential notation

Real numbers can have a finite or infinite number of digits. In order to work with real numbers, the numbers have to be approximated using a finite number of digits - usually in a standard way to allow efficient recognition of its magnitude. This is the reason why **scientific notation** is used, leading to the establishment of IEEE-754 for use in computing.

Due to the fundamental design of computers and the finite number of digits used in calculations, there will always be an inherent difference between the calculated value and the real value.

1.2 Error analysis

Any approximation of a function necessarily allows a possibility of deviation from the correct value of the function.

Error: The term used to denote the amount by which an approximation fails to equal the exact solution.

There are various kinds of errors:

- **Absolute error**: Value when $\hat{x}$ used instead of x

$$|\hat{x} - x|$$

- **Relative error**: Value when $\hat{x}$ used instead of x

$$\frac{|\hat{x} - x|}{|x|}$$

There are multiple possible sources of error:

- **Measurement error**: Self-explanatory.
- **Truncation error**: Occurs when a calculated number has more digits than available and some of them must be "forgotten".
- **Rounding error**: Occurs when a number is rounded to a specified precision.
- **Loss of significance** or **Cancellation error**: Occurs when an operation on two numbers increases relative error substantially more than the absolute error.

2 Ordinary Differential Equations (ODE)

The main application of numerical methods is for integrating differential equations (which why it is called Numerical Methods - for “solving them numerically”).

2.1 Basics of an ODE

Ordinary Differential Equation (ODE): An equality $A = B$ in which the only unknown (y) is a function of one variable whose derivative of some order appears explicitly (x) or (t).

The *ordinary* in ODE is the condition on the unknown being a function of a **single variable** i.e. there are no partial derivatives.

A typical ODE has a general form of:

$$y' = f(x, y)$$

and **may** have an exact (analytical) solution. However, more complicated functions $f(x, y)$ make it difficult and unsustainable to find. Numerical methods are faster and more efficient. These methods also apply to higher order ODEs such as 2nd Order Differential Equations.

2.2 Euler's Method

Euler's method is a *first-order* numerical method for solving ODEs with a given initial value. It is the most basic explicit method for numerical integration of ordinary differential equations and is the simplest Runge–Kutta method.

The Euler method is a first-order method, which means that the local error (error per step) is proportional to the square of the step size, and the global error (error at a given time) is proportional to the step size. It forms the basis for more complex methods such as the "predictor-corrector method".

Euler's method proposes a method to approximate the solutions to differential equations through the idea of **local linearity** or **linear approximation**, where small tangent lines over a short distance is used to approximate the solution to an initial-value problem.

Given $f(t, y)$ is a known function and values in the initial condition are also known. If y is **continuous** over some interval, there will be a **unique solution** to the ODE which can be approximated.

2.2.1 The intuition behind Euler's method

The aim of Euler's method is to **approximate** a function near a given dependent variable x or t e.g. $t = t_0$:

$$\frac{dy}{dt} = f(t, y) \text{ where } y(t_0) = y_0$$

- There are two known pieces of information derived from the initial condition and derivative function:

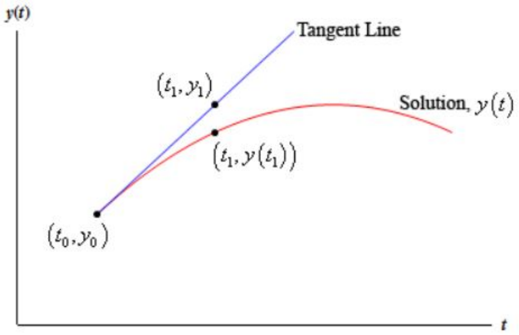
  1. Value of the function at $t = t_0$ from the **initial condition**.

  2. Value of the derivative at $t = t_0$ from plugging in t_0 into the given $\frac{dy}{dt}$.

- The information allows the equation of the tangent line to be calculated at $t = t_0$:

$$
\begin{aligned}
&y = mx + c \\
\Rightarrow\ &y = f(t_0, y_0)(t - t_0) + y_0
\end{aligned}
$$

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



- If t_1 is close enough to t_0 then the point y_1 on the tangent line is a good approximation to the actual value of the solution $y(t_1)$ at t_1.

  y_1 is found by plugging t_1 into the equation for the tangent line:

$$y_1 = y_0 + f(t_0, y_0)(t_1 - t_0)$$

- By proceeding in a similar manner, it could be repeated to get the values on the function. But only the value t_0 is known and $t_1, t_2, \ldots$ is unknown. However, recall that y_1 **is an approximation** to t_1 and, if it a good approximation, can be used as t_1 to obtain t_2:

$$y_2 = y_1 + f(t_1, y_1)(t_2 - t_1)$$

- The process is then repeated:

$$
\begin{aligned}
&y_3 = y_2 + f(t_2, y_2)(t_3 - t_2) \\
&y_4 = y_3 + f(t_3, y_3)(t_4 - t_3) \\
&\ldots
\end{aligned}
$$

The general equation of Euler's method is:

$$y_{n+1} = y_n + f_n \cdot (t_{n+1} - t_n)$$

It is common to simplify the **step size** $t_{n+1} - t_n$ into h:

$$y_{n+1} = y_n + h\, f_n$$

Pseudocode:

Listing 1: Euler's Method: Pseudocode

```
Start
    Define function f(x,y)

    Read values of initial condition(x0 and y0),
       number of steps (n) and calculation point (xn)

    Calculate step size (h) = (xn - x0)/n

    Set i=0

    Loop
        yn = y0 + h *  f(x0 + i*h, y0)

        y0 = yn

        i = i + 1
    while i < n

    Display yn as result
Stop
```

Python:

Listing 2: Euler's Method: Python code

```python
"""
Euler's method.py:

The classical Euler's method to solve dy\dx = -6x
"""

import numpy as np
import matplotlib.pyplot as plt

x0 = 0
y0 = 1
h = 0.03
x_final = 1
N = round((x_final-x0)/h)
```

```
x = np.linspace(0,x_final,N)
y = np.zeros(len(x))
y[0] = y0

for n in range(0,len(x) -1):
    y[n+1] = (y[n] + h*(-6*y[n])) # Input the
        function equation

y_true = np.exp(-6*x) # Plot the true graph of
    function 'exp(-6x)'

""" Plotting program """

plt.plot(x,y,'b.-',x,y_true,'r-')
plt.legend(['Euler','True'])
plt.axis([0,2,0,9])
plt.grid(True)
plt.title("Solution of $y'=y , y(0)=1$")
plt.show()
```

2.2.2 Error analysis

In Euler's method error analysis, the **local error** and **global error** is of concern.

The principal tool in the determination of truncation error is Taylor's theorem, a key result in calculus that provides a formula for the amount of error in using a truncated Taylor series to represent the corresponding function.

Convention: Exact solution to ODE denoted by lower-case letter. Numerical approximations to ODE denoted by capital letter.

Local truncation error: A measure of accuracy over one step of a method for the numerical solution of ODEs.

The *local error* or *error per step* is **proportional to the square of the step size** h.

Derivation: Taylor series for a function $x(t)$ defined as:

$$x(t+h) = x(t) + h\frac{dx}{dt}(t) + \frac{h^2}{2!}\frac{d^2x}{dt^2}(t) + \frac{h^3}{3!}\frac{d^3x}{dt^3}(t) + ...$$

or

$$x(t+h) = [x(t)] + h[x'(t)] + \frac{h^2}{2!}[x''(t)] + \frac{h^3}{3!}[x'''(t)] + ...$$

Euler's method can be seen as using Taylor series **truncated after the second term** $h[x'(t)]$ and given the value of $x(t)$ with its derivatives of x at t.

It is possible to find the value of $x(t+h)$ for any given h by choosing a small enough h - it will provide a good approximation to the value of $x(t+h)$.

$$X(t+h) = x(t) + h\frac{dx}{dt}(t) = x(t) + hf(t,x)$$

For example, applying it to points given the starting point (t_0, x_0):

$$\begin{aligned} X(t_1) &= X(t_0 + h) = x(t_0) + hf(t_0, x_0) \\ X(t_2) &= X(t_1 + h) = X(t_1) + hf(t_1, X_1) \\ &\dots \end{aligned}$$

Analysing Euler's method through application of Taylor series allows the accuracy of Euler's method to be evaluated and to obtain more accurate methods.

As stated earlier on the Taylor series:

$$x(t+h) = [x(t)] + h[f(t,x)] + \frac{h^2}{2!}[f'(t,x)]$$

Using the **order notation**, the equation is abbreviated to:

$$x(t+h) = [x(t)] + h[f(t,x)] + O(h^2)$$

where $\frac{h^2}{2}f'(t,x) = O(h^2)$

Combining it with Euler's method $X(t+h) = x(t) + hf(t,x)$:

$$X(t+h) = x(t+h) + O(h^2)$$

The equation indicates that there is an **local truncation error** of order h^2 occurring in each loop from $x(t)$ to $x(t+h)$. Some errors may cancel out as some cases it will be below true value and some cases it will be above true value.

This is the reason Euler's method is known as a **first-order method**. The error, as it is dependent on the size of h, decreases as h decreases i.e. halving the size of h decreases error by four times.

Global truncation error: The error in the value of $X(t_0 + a)$ obtained using a numerical method to advance the required number of steps from a known value of $x(t_0)$

The *global error* or *error at a given time* is proportional to the step size - $O(h)$.

Derivation:

- Given a starting point (x_0, y_0) and using Euler's method with a step size of h, a value of $X(t_0 + 4)$ will require $\frac{4}{h}$ steps.
- Total error in $X(t_0 + 4)$ will be the sum of errors at each each step - $\frac{4}{h}$ times the value of typical step error.

- Total error is order of $\left[\frac{4}{h}\right]O(h^2)$ which is $O(h)$ i.e. halving step size will halve the error in the solution.

2.2.3 Predictor-corrector method

Improved Euler's method (Heun's method) is a second-order method and is composed of two equations:

1. Predictor equation
2. Corrector equation

Predictor equation:

In Euler's method, the gradient **at the beginning of the interval** $[x_i, x_{i+1}]$:

$$y'_i = f(x_i, y_i)$$

is used to make a prediction of the value of y_{i+1} at the end of the interval:

$$y^p_{i+1} = y_i + hf(x_i, y_i)$$

where p as superscript indicates a prediction.

Corrector equation:

Using the predicted value, an estimate of the gradient **at the end of the interval** is found:

$$y'_{i+1} = f(x_{i+1}, y^p_{i+1})$$

With the two gradient estimates at each end of the interval, an average can be found which is a better estimate on the gradient over range $[x_i, x_{i+1}]$:

$$y'_{av} = \frac{y'_i + y'_{i+1}}{2} = \frac{f(x_i, y_i) + f(x_{i+1}, y^p_{i+1})}{2}$$

Using the average gradient and Euler's method, the corrector equation is defined:

$$y_{i+1} = y_i + h\left(\frac{y'_i + y'_{i+1}}{2}\right) = y_i + h\left(\frac{f(x_i, y_i) + f(x_{i+1}, y^p_{i+1})}{2}\right)$$

Improved Euler's method:

- Predictor equation:

$$y^p_{i+1} = y_i + h[f(x_i, y_i)]$$

- Corrector equation:

$$y_{i+1} = y_i + h\left[\frac{f(x_i, y_i) + f(x_{i+1}, y^p_{i+1})}{2}\right]$$

Example 1: Solve the problem: $\frac{dx}{dt} = \frac{x^2}{t+1}$ given $x(0) = 1$ with Improved Euler's method.

1. Predictor equation:

$$\begin{aligned}\hat{X}_1 &= x_0 + h[f(t_0, x_0)] \\ &= x_0 + h\left[\frac{x^2}{t+1}\right] \\ &= 1 + 0.1\frac{1^2}{0+1} \\ &= 1.100\end{aligned}$$

2. Corrector equation:

$$\begin{aligned}X_1 &= x_0 + \frac{1}{2}h[f(t_0, x_0) + f(t_1, \hat{X}_1)] \\ &= x_0 + \frac{1}{2}h\left[\frac{x_0^2}{t_0+1} + \frac{\hat{X}_1^2}{t_1+1}\right] \\ &= 1.000 + \frac{1}{2}0.1\left[\frac{1^2}{0+1} + \frac{1.100^2}{0.100+1}\right] \\ &= 1.105\end{aligned}$$

2.3 Runge-Kutta 2^{nd} order methods

Runge-Kutta 2^{nd} Order method is a numerical technique used to solve a 1^{st} ordinary differential equation of the form:

$$\frac{dy}{dx} = f(x, y) \quad y(0) = y_0$$

Example 1: Rewrite $\frac{dy}{dx} + 2y = 1.3e^{-x} \quad y(0) = 5$ in the above format.

$$\frac{dy}{dx} = 1.3e^{-x} - 2y \quad y(0) = 5$$

In the previous section with Euler's method, it was derived using information on the slope and the derivative of y at the given time step to derive the solution to the next time step. Runge-Kutta methods are a **class of methods** that uses the same information at more than one point to find the next time step.

Euler's method was defined as:

$$y_{i+1} = y_i + f(x_i, y_i)h$$

where: $x_0 = 0$, $y_0 = y(x_0)$ and $h = x_{i+1} - x_i$

Deriving Euler's method from Taylor series:

$$\begin{aligned}y_{i+1} &= y_i + \frac{dy}{dx}|_{x_i,y_i}(x_{i+1} - x_i) + \frac{1}{2!}\frac{d^2y}{dx^2}|_{x_i,y_i}(x_{i+1} - x_i)^2 + \frac{1}{3!}\frac{d^3y}{dx^3}|_{x_i,y_i}(x_{i+1} - x_i)^3 + ... \\ &= y_i + f(x_i, y_i)(x_{i+1} - x_i) + \frac{1}{2!}f'(x_i, y_i)(x_{i+1} - x_i)^2 + \frac{1}{3!}f''(x_i, y_i)(x_{i+1} - x_i)^3\end{aligned}$$

Notice the first two terms of the Taylor series: $y_{i+1} = y_i + f(x_i, y_i)h$ are Euler's method and thus considered to be Runge-Kutta 1^{st} order method. The Runge-Kutta 2^{nd} order method will include one more term of the Taylor series:

$$y_{i+1} = y_i + f(x_i, y_i)(x_{i+1} - x_i) + \frac{1}{2!}f'(x_i, y_i)(x_{i+1} - x_i)^2$$

However, there is certain difficulties in finding $f'(x, y)$. Runge and Kutta simplified it and wrote the general Runge-Kutta expression as:

$$\boxed{y_{i+1} = y_i + h\phi(x_i, y_i, h)}$$

where:

- ϕ is the **increment function**: $\phi = a_1k_1 + a_2k_2 + ... + a_nk_n$
- a_i are constants
- k_i:

$$\begin{aligned}
&k_1 = f(x_i, \quad y_i)\\
&k_2 = f(x_i + p_1h, \quad y_i + q_{11}k_1h)\\
&k_3 = f(x_i + p_2h, \quad y_i + q_{21}k_1h + q_{22}k_2h)\\
&\;\;.\\
&\;\;.\\
&\;\;.\\
&k_n = f(x_i + p_{n-1}h, \quad y_i + q_{n-1,1}k_1h + q_{n-1,2}k + 2h + ... + q_{n-1,n-1}k_{n-1}h)
\end{aligned}$$

Example 2: 2^{nd} order method is defined as:

$$y_{i+1} = y_i + (a_1k_1 + a_2k_2)h$$

where:

- $k_1 = f(x_i, \quad y_i)$
- $k_2 = f(x_i + p_1h, \quad y_i + q_{11}k_1h)$

2.3.1 Derivation of Second-order Runge-Kutta

The value of a, b, p and q need to be found.

1. Being with the equation:

$$y_{i+1} = y_i + (a_1k_1 + a_2k_2)h$$

   where:

   - $k_1 = f(x_i, \quad y_i)$

- $k_2 = f(x_i + p_1 h, \quad y_i + q_{11} k_1 h)$

2. Use the Taylor expansion:

$$y(x+h) = y(x) + hy'(x) + \frac{h^2}{2!}y''(x) + \dots$$

simplifies to:

$$y_{i+1} = y_i + hf(x_i, y_i) + \frac{h^2}{2}f'(x_i, y_i) + \dots$$

3. Use partial differentiation:

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

4. Substitute into the Taylor expansion:

$$y_{i+1} = y_i + hf(x_i, y_i) + \frac{h^2}{2}\left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx}\right) + \dots$$

5. Recall Taylor expansion for function with two variables:

$$F(x+h, \quad y+k) = F(x,y) + h\frac{\partial F}{\partial x} + k\frac{\partial F}{\partial y} + O(h^2, k^2)$$

6. Consider $k_2 = f(x_i + ph, \quad y_i + qk_1 h)$ and using Taylor expansion above:

$$\begin{aligned} k_2 &= f(x_i + ph, \quad y_i + qk_1 h) \\ &= f(x_i, y_i) + ph\frac{\partial f}{\partial x} + qk_1 h\frac{\partial f}{\partial y} + O(h^2) \end{aligned}$$

7. Substitute for k_1 and k_2 into $y_{i+1} = y_i + h(ak_1 + bk_2)$:

$$\begin{aligned} y_{i+1} &= y_i hak_1 + hbk_2 \\ &= y_i haf(x_i, y_i) + hb\left[f(x_i, y_i) + ph\frac{\partial f}{\partial x} + qf(x_i, y_i)h\frac{\partial f}{\partial y} + O(h^2)\right] \\ &= y_i + (a+b)hf(x_i, y_i) + h^2\left(bp\frac{\partial f}{\partial x} + bq\frac{\partial f}{\partial y}f(x_i, y_i)\right) + O(h^3) \end{aligned}$$

8. Compare the two equations:

$$y_{i+1} = y_i + hf(x_i, y_i) + \frac{h^2}{2}\left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx}\right) + \dots$$

and

$$y_{i+1} = y_i + (a+b)hf(x_i, y_i) + h^2\left(bp\frac{\partial f}{\partial x} + bq\frac{\partial f}{\partial y}f(x_i, y_i)\right) + O(h^3)$$

9. Notice possible value:

   - $a + b = 1$
   - $bp = \frac{1}{2}$
   - $bq = \frac{1}{2}$

Example 3: Heun's Method: Given $b = \frac{1}{2}$, $a = \frac{1}{2}$ and $p = q = 1$:

$$y_{i+1} = y_i + h\left(\frac{1}{2}k_1 + \frac{1}{2}k_2\right)$$

where the stages of Runge-Kutta $k_1 = f(x_1, y_1)$ and $k_2 = f(x_i + h, \quad y_i + k_1 h)$

There are infinitely many types of Runge-Kutta methods due to the fact that each choice of b gives a new variation with roughly the same global error $O(h^2)$ with a few minor differences.

2.3.2 Improved Euler - Midpoint Method

The Midpoint Method is another method that improved Euler's method that is based on using Euler's method to estimate the gradient **at the midpoint of the interval** $[x_i, x_{i+1}]$. This gives a better estimate of y at the endpoint.

Midpoint method:

$$y_0 = y(x_0)$$
$$y_{i+\frac{1}{2}} = y_i + \frac{h}{2}f(x_i, y_i)$$
$$y_{i+1} = y_i + hf(x_{i+\frac{1}{2}}, y_{i+\frac{1}{2}})$$

Operation:

$$k_1 = f(x_i, y_i)$$

- Using Euler's method to calculate gradient at x_i:

$$k_2 = f(x_i + \frac{1}{2}h, \quad y_i + \frac{1}{2}hk_1)$$

- Using Euler's method to calculate gradient at $x_i + \frac{h}{2}$:

$$k_2 = f(x_i + \frac{1}{2}h, \quad y_i + \frac{1}{2}hk_1)$$

- Value used to calculate the endpoint:

$$y_{i+1} = y_i + hk_2$$

Local truncation error of order $O(h^3)$ and **Global truncation error** of $O(h^2)$.

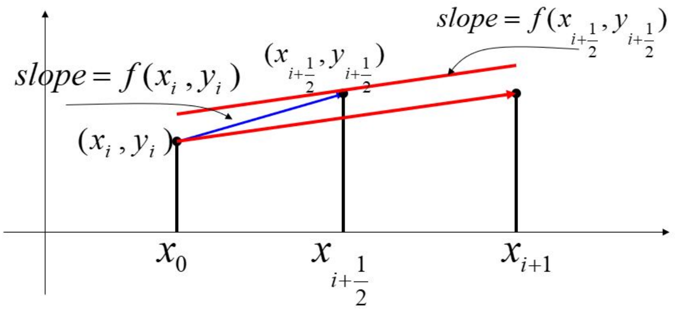

$slope = f(x_{i+\frac{1}{2}}, y_{i+\frac{1}{2}})$
$(x_{i+\frac{1}{2}}, y_{i+\frac{1}{2}})$
$slope = f(x_i, y_i)$
(x_i, y_i)
x_0
$x_{i+\frac{1}{2}}$
x_{i+1}

Figure 1: Graphical representation of Midpoint Method

2.3.3 Higher-order Runge-Kutta methods

There are infinitely many types of higher-order Runge-Kutta methods for example the classic 4^{th} order Runge-Kutta method:

$$\begin{aligned}
k_1 &= f(x_i, \quad y_i) \\
k_2 &= f(x_i + \frac{1}{2}h, \quad y_i + \frac{1}{2}k_1 h) \\
k_3 &= f(x_i + \frac{1}{2}h, \quad y_i + \frac{1}{2}k_2 h \\
k_4 &= f(x_i + h, \quad y_i + k_3 h)
\end{aligned}$$

3 Higher order ODE - Coupled 1^{st} order ODE

The classes of second- and higher order differential equations with no analytical solution can be solved by numerical means. Numerical solutions of second- and higher-order equations does not need any new theory or technique.

3.1 Numerical solution of coupled first-order equations

Given that first-order differential equations involve a single dependent variable and a single independent variable, higher-order differential equations can be broken down into sets of coupled first-order equations. Each equation involve the **same** independent variable but with **more than one** dependent variable such as:

$$\begin{aligned}
\frac{dx}{dt} &= x - y^2 + xt = f_1(t, x, y) \\
\frac{dy}{dt} &= 2x^2 + xy - t = f_2(t, x, y)
\end{aligned}$$

The derivative of each dependent variables x and y depends on the independent variable t and the other dependent variable.

Original Euler's method:

$$X_{n+1} = X_n + hf(t_n, X_n)$$

Euler's method on coupled first-order equations:

$$X_{n+1} = X_n + hf_1(t_n, X_n, Y_n)$$

meaning value of X_{n+1} depends on: t_n, X_n and Y_n.

Example 1: Using Euler's method with time step $h = 0.1$, find the value of $X(1.2)$ satisfying the following initial value problem:

$$\frac{dx}{dt} = x - y^2 + xt, \quad x(1) = 0.5$$

$$\frac{dy}{dt} = 2x^2 + xy - t, \quad y(1) = 1.2$$

1. Denote right-hand side of two equations as $f_1(t, x, y)$ and $f_2(t, x, y)$, first step is:

$$f_1(t, x, y) = x - y^2 + xt \quad \text{and} \quad f_2(t, x, y) = 2x^2 + xy - t$$

2. Note imposed initial equations $x_0 = 0.5$ and $y_0 = 1.2$:

$$\begin{aligned} X_1 &= x_0 + hf_1(t_0, x_0, y_0) \\ &= 0.5 + 0.1f_1(1, 0.5, 1.2) \\ &= 0.4560 \end{aligned} \qquad \begin{aligned} Y_1 &= y_0 + hf_2(t_0, x_0, y_0) \\ &= 1.2 + 0.1f_2(1, 0.5, 1.2) \\ &= 1.2100 \end{aligned}$$

3. Second step is (final step since $1 + 2(0.1) = 1.2$):

$$\begin{aligned} X_2 &= X_1 + hf_1(t_1, x_1, y_1) \\ &= 0.4560 + 0.1f_1(1.1, 0.4560, 1.2100) \\ &= \underline{0.4054} \end{aligned} \qquad \begin{aligned} Y_2 &= Y_1 + hf_2(t_1, x_1, y_1) \\ &= 1.2100 + 0.1f_2(1.1, 0.4560, 1.2100) \\ &= 1.1968 \end{aligned}$$

Example 2: Find the value of $X(1.2)$ satisfying the initial-value problem:

$$\frac{dx}{dt} = x - y^2 + xt \quad x(1) = 0.5$$

$$\frac{dy}{dt} = 2x^2 + xy - t \quad y(1) = 1.2$$

using the second-order predictor-corrector method with time step $h = 0.1$.

- First step:

  1. Find the predictor equation, noting the initial conditions:

$$\begin{aligned} \hat{X}_1 &= x_0 + hf_1(t_0, x_0, y_0) \\ &= 0.4560 \end{aligned} \qquad \begin{aligned} \hat{Y}_1 &= y_0 + hf_2(t_0, x_0, y_0) \\ &= 1.2100 \end{aligned}$$

2. Find the corrector equation, using the values obtained from the predictor:

$$\begin{aligned}\hat{X}_1 &= x_0 + \frac{1}{2}h[f_1(t_0, x_0, y_0) + f_1(t_1, \hat{X}_1, \hat{Y}_1)] \\ &= 0.5 + 0.05[f_1(1, 0.5, 1.2) + f_1(1.1, 0.456, 1.21)] \\ &= 0.4527\end{aligned}$$

$$\begin{aligned}\hat{Y}_1 &= y_0 + \frac{1}{2}h[f_2(t_0, x_0, y_0) + f_2(t_1, \hat{X}_1, \hat{Y}_1)] \\ &= 1.2 + 0.05[f_2(1, 0.5, 1.2) + f_2(1.1, 0.456, 1.21)] \\ &= 1.1984\end{aligned}$$

- Second step:

1. Find the predictor equation, noting the initial conditions:

$$\begin{aligned}\hat{X}_2 &= x_1 + hf_1(t_1, x_1, y_1) \qquad & \hat{Y}_2 &= y_1 + hf_2(t_1, x_1, y_1) \\ &= 0.4042 & &= 1.1836\end{aligned}$$

2. Find the corrector equation, using the values obtained from the predictor:

$$\begin{aligned}\hat{X}_2 &= X_1 + \frac{1}{2}h[f_1(t_1, X_1, Y_1) + f_1(t_2, \hat{X}_2, \hat{Y}_2)] \\ &= 0.4527 + 0.05[f_1(1.1, 0.4527, 1.1984) + f_1(1.2, 0.4042, 1.1836)] \\ &= \underline{0.4028}\end{aligned}$$

$$\begin{aligned}\hat{Y}_2 &= Y_1 + \frac{1}{2}h[f_2(t_1, Y_1, Y_1) + f_2(t_2, \hat{X}_2, \hat{Y}_2)] \\ &= 1.1984 + 0.05[f_2(1.1, 0.4527, 1.1984) + f_2(1.2, 0.4042, 1.1836)] \\ &= 1.1713\end{aligned}$$

4 Finite Differences for ODEs (Boundary value problems)

First-order ODEs have only one boundary condition treated as an initial condition. With second- and higher-order, there are more than one boundary conditions such as:

$$y'' - 2y' + 3y = \sin(x) \quad \text{with} \quad y(0) = 0 \text{ and } y(1) = 0$$

Boundary value problems do not always have solutions and prior analysis would usually be required to determine the presence of a solution. For Year 2 Numerical Methods, boundary value problems are assumed to have a solution.

The previous approach will not work as the boundary conditions provided do not provide enough information to approximate it using a system of first-order equations and there is no guarantee the iteration would lead to the solution as given.

4.1 Central difference method

One method of solving boundary value problems is based in by breaking up the interval in question.

An important concept is **Central Difference**:

$$\delta y(x_i) = y\left(x_i + \frac{h}{2}\right) - y\left(x_i - \frac{h}{2}\right)$$

Also note the approximation of derivative $\frac{dy}{dx}|_{x=x_i}$:

$$\frac{1}{h}\delta y(x_i) = \frac{y\left(x_i + \frac{h}{2}\right) - y\left(x_i - \frac{h}{2}\right)}{h} \approx \frac{dy}{dx}|_{x=x_i}$$

Derivation:

1. Define first central difference:

$$\delta y(x_i) = y\left(x_i + \frac{h}{2}\right) - y\left(x_i - \frac{h}{2}\right)$$

2. Define second central difference $\delta^2 y(x_i)$:

$$\begin{aligned}\delta^2 y(x_i) &= \delta[\delta y(x_i)]\\ &= \delta y\left(x_i + \frac{h}{2}\right) - \delta y\left(x_i - \frac{h}{2}\right)\\ &= \left\{y\left[\left(x_i + \frac{h}{2} + \frac{h}{2}\right)\right] - \left[\left(x_i + \frac{h}{2}\right) - \frac{h}{2}\right]\right\} - \left\{y\left[\left(x_i - \frac{h}{2}\right) + \frac{h}{2}\right] - y\left[\left(x_i - \frac{h}{2}\right) - \frac{h}{2}\right]\right\}\\ &= y(x_i + h) - 2y(x_i) + y(x_i - h)\end{aligned}$$

3. Second derivative y'' is defined as:

$$\frac{1}{h^2}\delta^2 y(x_i) \approx \frac{d^2 y}{dx^2}|_{x=x_i}$$

4. Approximate of first derivative defined as:

$$\begin{aligned}\frac{1}{2h}\Delta y(x_i) &= \frac{y(x_i + h) - y(x_i - h)}{2h}\\ &= \frac{y(x_{i+1}) - y(x_{i-1})}{2h}\\ &= \frac{dy}{dx}|_{x=x_i}\end{aligned}$$

5. Let $u_i \approx y(x_i)$

Central difference method:

$$\frac{dy}{dx}|_{x=x_i} \approx \frac{u_{i+1} - u_{i-1}}{2h} \quad \text{and} \quad \frac{d^2y}{dx^2}|_{x=x_i} \approx \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2}$$

Example 1: Solve the given model convection-diffusion equation:

$$y'' - 20y' = -1 \quad \text{where} \quad y(0) = 1 \text{ and } y(1) = 0$$

1. Divide interval of interest $[0, 1]$ into five segments, forming step size $h = 0.2$ into the following: $x_0 = 0$, $x_1 = 0.2$, $x_2 = 0.4$, $x_3 = 0.6$, $x_4 = 0.8$ and $x_5 = 1.0$.

2. Estimate derivatives using formulae:

   $$\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} - 20\left(\frac{u_{i+1} - u_{i-1}}{2h}\right) = -1$$

   Rearrange into:

   $$\left(\frac{10}{h} + \frac{1}{h^2}\right)u_{i-1} - \frac{2}{h^2}u_i + \left(-\frac{10}{h} + \frac{1}{h^2}\right)u_{i+1} = -1$$

   $$au_{i-1} + bu_i + cu_{i+1} = -1$$

3. For each i, there is an equation defined by u_{i-1}, u_i and u_{i+1}:

   $$\begin{aligned} i = 1: &\quad au_0 + bu_1 + cu_2 = -1 \\ i = 2: &\quad au_1 + bu_2 + cu_2 = -1 \\ i = 3: &\quad au_2 + bu_3 + cu_4 = -1 \\ i = 4: &\quad au_3 + bu_4 + cu_5 = -1 \end{aligned}$$

   where u_0 and u_5 are known boundary values.

4. Express in matrix form:

   $$\begin{bmatrix} b & c & 0 & 0 \\ a & b & c & 0 \\ 0 & a & b & c \\ 0 & 0 & a & b \end{bmatrix} * \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} -1 - au_0 \\ -1 \\ -1 \\ -1 - cu_5 \end{bmatrix}$$

5. Use Gaussian elimination to obtain the solutions:

   $$u_1 \approx 0.9928 \quad u_2 \approx 1.0544 \quad u_3 \approx 0.9095 \quad u_4 \approx 1.3843$$

5 Partial Differential Equations

The method of finite differences extends to solution of partial differential equations. Many solutions exists for many types of partial differentials. Two examples of Partial Differential Equations and a method to solve are given.

5.1 Heat/Diffusion equation

General form of the Heat/Diffusion equation:

$$\frac{\partial y}{\partial t} = \frac{\partial^2 y}{\partial x^2} \quad 0 < x < 1 \text{ and } t > 0$$

with boundary conditions $y(0,t) = y(1,t) = 0$ and initial condition $y(x,0) = y_0(x)$

The heat/diffusion equation describes the temperature distribution in a thin metal rod of unit length with insulated sides and both ends kept at $0K$.

At time $t = 0$, distribution is $y_0(x)$ and solution to the partial differential describes evolution in time. It can also describe other diffusive processes such as pollutant gas in a pipe or random walk.

Finite differences (central difference) will be used for both derivatives:

1. Divide the spatial interval into N segments of equal length $h = \frac{1}{N}$:

$$x_0 = 0; x_1 = h; x_2 = 2h; \ldots ; x_N = Nh = 1$$

2. Divide the time domain into equal segments of length k:

$$t_0 = 0; t_1 = k; t_2 = 2k$$

where there is no clear end, the process is repeated as far as required

3. Let U_j^m be value $y(x_j, t_m)$ where the initial condition gives the values of y at $t = 0$:

$$U_j^0 = y(x_j, t_0) = y(x_j, 0) = y_0(x_j) \quad \text{for } j = 0, 1, 2, \ldots, N$$

4. The boundary conditions gives values for $j = 0$ corresponding to $x =$ and for $j = n$ corresponding to $x = 1$:

$$U_j^m = y(x_j, t_m) \to y(0, t_m) = U_0^m = 0 \quad \text{and} \quad y(1, t_m) = U_N^m = 0$$

where $\forall m \geq 0$

5. Recall function of two variables to deal with spatial derivative:

$$\frac{\partial^2 y(x,t)}{\partial x^2} \approx \frac{y(x+h,t) - 2y(x,t) + y(x-h,t)}{h^2}$$

where t remains constant and x varies by h in both directions

Working example 1: Given $t = t_m$ and $x = x_j$ which leads to $x \pm h = x_{j\pm1}$

$$\begin{aligned} \frac{\partial^2 y(x_j, t_m)}{\partial x^2} &\approx \frac{y(x_{j+1}, t_m) - 2y(x_j, t_m) + y(x_{j-1}, t_m)}{h^2} \\ &\approx \frac{U_{j+1}^m - 2U_j^m + U_{j-1}^m}{h^2} \quad j = 1, 2, \ldots, N-1 \end{aligned}$$

6. For the time derivative, use a **"forward difference"**:

$$\frac{\partial y(x,t)}{\partial t} \approx \frac{y(x,t+k)-y(x,t)}{k}$$

A forward difference must be used as there cannot be $t=-1$

Working example 2: Given $t=t_m$ and $x=x_j$ which leads to $t+k=t_{m+1}$

$$\begin{aligned}\frac{\partial y(x_j,t_m)}{\partial t} &\approx \frac{y(x_j,t_{m+1})-y(x_j,t_m)}{k}\\ &\approx \frac{U_j^{m+1}-U_j^m}{k} \qquad m=0,1,2,3\ldots\end{aligned}$$

7. Set two derivatives equal:

$$\frac{U_{j+1}^m-2U_j^m+U_{j-1}^m}{h^2}=\frac{U_j^{m+1}-U_j^m}{k}$$

Let $v=\frac{k}{h^2}$

$$v\left(U_{j+1}^m-2U_j^m+U_{j-1}^m\right)=U_j^{m+1}-U_j^m$$

8. Express in terms of U_j^{m+1}:

$$\begin{aligned}U_j^{m+1}&=U_j^m+v\left(U_{j+1}^m-2U_j^m+U_{j-1}^m\right)\\ &=vU_{j-1}^m+(1-2v)U_j^m+vU_{j+1}^m\end{aligned}$$

5.2 Laplace/Poisson equation

Laplace's second-order equation is widely used in physics since its solutions occur in fields such as electrical, magnetic and gravitational. Poisson's equation is a generalisation of Laplace's equation that is widely used in theoretical physics e.g. the solution to the potential field caused by a given electrical charge or mass density distribution.

Poisson' equation:

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=g(x,y)$$

where Laplace's equation is given when $g=0$

The central difference is used to approximate the two derivatives:

$$\begin{aligned}\frac{\partial^2 u(x,y)}{\partial x^2} &\approx \frac{u(x+h,y)-2u(x,y)+u(x-h,y)}{h^2}\\ \frac{\partial^2 u(x_i,y_j)}{\partial x^2} &\approx \frac{u(x_{i+1},y_j)-2u(x_i,y_j)+u(x_{i-1},y_j)}{h^2}\\ &=\frac{U_{i+1}^j-2U_i^j+U_{i-1}^j}{h^2}\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial^2 u(x,y)}{\partial y^2} &\approx \frac{u(x,y+h)-2u(x,y)+u(x,y-h)}{h^2} \\
\frac{\partial^2 u(x_i,y_j)}{\partial y^2} &\approx \frac{u(x_i,y_{j+1})-2u(x_i,y_j)+u(x_i,y_{j-1})}{h^2} \\
&= \frac{U_i^{j+1}-2U_i^j+U_i^{j-1}}{h^2}
\end{aligned}$$

As mentioned earlier, Laplace's equation is found when the sum of the two derivatives is set to 0. An approximation of Laplace's equation is found when using the central difference of the two derivatives:

$$\frac{U_{i+1}^j - 2U_i^j + U_{i-1}^j}{h^2} + \frac{U_i^{j+1} - 2U_i^j + U_i^{j-1}}{h^2} = 0$$

Simplifies to:

$$U_{i+1}^j + U_{i-1}^j + U_i^{j+1} + U_i^{j-1} - 4U_i^{j-1} - 4U_i^j = 0$$

Worked Example 1: The simplest case of Laplace's equation is on a rectangle: $0 \leq x \leq a$ and $0 \leq y \leq b$ divided into $(n-1)x(m-1)$ squares with side length h. This means:

- $a = nh$ and $b = mh$
- $i = 1 \dots n$ and $j = 1 \dots m$

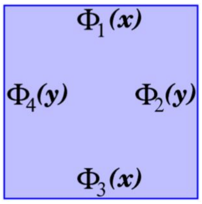


Figure 2: Graphical representation of the locations of ϕ

The solution requires knowledge of u on the shape boundaries (Boundary Conditions):

- $\phi_1(x)$ where $y = b$ giving the values U_i^m for $i = 1 \dots n$
- $\phi_2(x)$ where $x = a$ giving the values U_n^j for $i = 1 \dots m$
- $\phi_3(x)$ where $y = 0$ giving the values U_i^1 for $i = 1 \dots n$
- $\phi_4(x)$ where $y = 0$ giving the values U_1^j for $i = 1 \dots m$

The values on the edges U_i^j are known from Boundary Conditions, only the interior points $P_1 \dots P_9$ are unknown. The approximation of Laplace's equation is defined:

$$U_{i+1}^j + U_{i-1}^j + U_i^{j+1} + U_i^{j-1} - 4U_i^j = 0$$

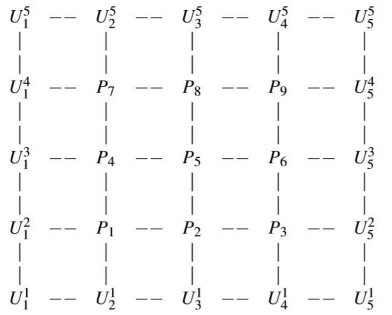

U_1^5 U_2^5 U_3^5 U_4^5 U_5^5
U_1^4 P_7 P_8 P_9 U_5^4
U_1^3 P_4 P_5 P_6 U_5^3
U_1^2 P_1 P_2 P_3 U_5^2
U_1^1 U_2^1 U_3^1 U_4^1 U_5^1

Figure 3: Square broken down into points with U known and P unknown

which is a relationship that involves five points:

- Center at x_i, y_i
- Above at U_i^{j+1}
- Below at U_i^{j-1}
- Right at U_{i+1}^j
- Left at U_{i-1}^j

This relationship can be applied where every unknown interior point is taken as the centre. For example:

$$\begin{aligned}
\text{at } P_1 &: U_1^2 + U_2^1 + P_2 + P_4 - 4P_1 = 0 \\
\text{at } P_2 &: P_1 + U_3^1 + P_3 + P_5 - 4P_2 = 0 \\
\text{at } P_3 &: P_2 + U_4^1 + U_5^2 + P_6 - 4P_3 = 0
\end{aligned}$$

The equations can be rewritten in matrix form to be solved as usual:

$$\begin{bmatrix}
-4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & -4 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & -4 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & -4 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & -4 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & -4 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & -4 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & -4 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & -4
\end{bmatrix} * \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \\ P_6 \\ P_7 \\ P_8 \\ P_9 \end{bmatrix}$$

As solving via matrix is a computationally expensive task, an alternative is the **Relaxation method**.

5.3 Relaxation method

The relaxation method is based on the **continuity of solutions**.

Continuity of solutions:
Able to take each point as the average of its four nearest neighbours:

$$\begin{aligned} U_j^i &= \frac{(U_{i+1}^j + U_{i-1}^j + U_i^{j+1} + U_i^{j-1})}{4} \\ &= \frac{U_{i+1}^j + U_{i-1}^j + U_i^{j+1} + U_i^{j-1} - 4U_i^j}{4} = 0 \end{aligned}$$

1. Set all points inside grid to an initial value k - average of values of known boundary points

2. For each interior point, calculate:

$$\left(U_i^j\right)_{new} = \left(U_i^j\right)_{old} + r_i^j$$

   where the **residual** r_{ij} is given by:

$$r_i^j = \frac{U_{i+1}^j + U_{i-1}^j + U_i^{j+1} + U_i^{j-1} - 4U_i^j}{4}$$

   - The residual vanishes as the averages are recalculated repeatedly

   - Set a desired accuracy $\epsilon > 0$ and stop when $|r_i^j| < \epsilon$ at every interior point

Poisson equation can be easily incorporated into the relaxation methods by taking $g(x_i, y_j) = g_i^j$ for example:

$$r_i^j = \frac{U_{i+1}^j + U_{i-1}^j + U_i^{j+1} + U_i^{j-1} - 4U_i^j - h^2 g_i^j}{4}$$

6 Summary

6.1 Numerical methods for initial-value problems

MIT Numerical

7 Introduction

7.1 Trapezoid Rule

Not every integral has an exact solution, the solution will have to be approximated using a numerical technique such as the Trapezoid Rule.

Given the integral of $f(x)$ over interval $[a, b]$, divide the interval into n segments of equal length defined as $h = \frac{b-a}{n}$.

From the picture:

- Set $a = x_0$ and $b = x_n$

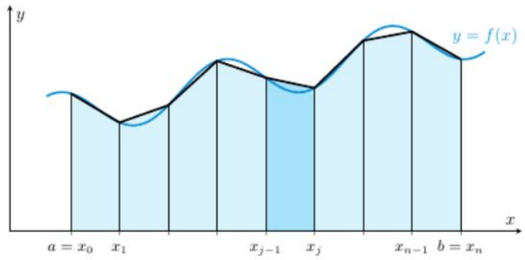


Figure 4: Graphical representation of the Trapezoid Rule

- n intervals are $[x_i, x_{i+1}]$ for $i = 0...n-1$

Since the area under the curve $f(x)$ is divided into segments, the area under the curve (integration):

$$\int_{x_i}^{x_{i+1}} f(x)\ \mathrm{dx}$$

is the approximated to the average of the area of left trapezoid $f(x_i)$ and right trapezoid $f(x_{i+1})$:

$$\int_{x_i}^{x_{i+1}} f(x)\ \mathrm{dx} \approx \frac{1}{2}h[f(x_i) + f(x_{i+1})]$$

The general case for the Trapezoid Rule is defined as:

$$\int_a^b f(x)\ \mathrm{dx} \approx \left(\frac{b-a}{2n}\right)\left[f(x_0) + 2\left(\sum_{i=1}^{n-1} f(x_i)\right) + f(x_n)\right]$$

Example 1: Given $n = 3$, $a = x_0$ and $b = x_3$, the general case becomes:

$$\frac{1}{2}f\left[f(x_0) + f(x_1)\right] + \frac{1}{2}h\left[f(x_1) + f(x_2)\right] + \frac{1}{2}f\left[f(x_2) + f(x_3)\right]$$

8 Change of Variables

8.1 One function of two random variables

Given two random variables X and Y characterized by the joint PDF $f_{X,Y}(x, y)$ and a function $g(x, y)$, a new random variable Z is formed as $Z = g(X, Y)$. The PDF of Z denoted $f_Z(z)$ can be calculated.

This situation occurs commonly in practical applications such as a receiver output signal Z which consists of the desired signal X that is buried in the noise Y thus leading to: $Z = X + Y$

If two random variables are independent then the density of their sum is equal to the convolution of their density function.

$$f_Z = f_X \otimes f_Y$$

$$f_Z(z) = \int_{x=-\infty}^{x=+\infty} f_{XY}(x, z-x)\, dx = \int_{y=-\infty}^{y=+\infty} f_{XY}(z-y, y)\, dy$$

If X and Y are independent then $f_{XY}(x,y) = f_X(x)f_Y(y)$:

$$f_Z(z) = \int_{y=-\infty}^{y=+\infty} f_X(z-y)f_Y(y)\, dy = \int_{x=-\infty}^{x=+\infty} f_X(x)f_Y(z-x)\, dx$$

8.2 Two functions of two random variables