Finite difference method for Linear boundary value problems: ①

Consider the linear boundary value problem

$$y'' = p(x)y' + q(x)y + r(x), \quad \text{for } a \le x \le b, \; y(a) = \alpha, \; y(b) = \beta.$$

Select an integer $N > 0$, divide $[a, b]$ into ~~equ~~ $(N+1)$ equal subintervals whose end points are the mesh points $x_i = a + ih$, $i = 0, 1, \ldots N+1$, where $h = \frac{b-a}{N+1}$.

The solution y is known at $x_0 = a$, and $x_{N+1} = b$.

At interior points, x_i, $i = 1, 2, \ldots N$, the differential equation is

$$y''(x_i) = p(x_i)y'(x_i) + q(x_i)y(x_i) + r(x_i). \quad \text{①}$$

Expand y using Taylor polynomial evaluated at x_{i+1} and x_{i-1}, we have by assuming $y \in C^4[x_{i-1}, x_{i+1}]$,

$$y(x_{i+1}) = y(x_i) + h y'(x_i) + \frac{h^2}{2} y''(x_i) + \frac{h^3}{6} y'''(x_i) + \frac{h^4}{24} y^{(4)}(\xi_i^+)$$

for some ξ_i^+ in (x_i, x_{i+1}), and

$$y(x_{i-1}) = y(x_i) - h y'(x_i) + \frac{h^2}{2} y''(x_i) - \frac{h^3}{6} y'''(x_i) + \frac{h^4}{24} y^{(4)}(\xi_i^-)$$

for some ξ_i^- in (x_{i-1}, x_i). By adding these two, we find

$$y(x_{i+1}) + y(x_{i-1}) = 2y(x_i) + h^2 y''(x_i) + \frac{h^2}{24}\left(y^{(4)}(\xi_i^+) + y^{(4)}(\xi_i^-)\right)$$

Therefore

$$y''(x_i) = \frac{y(x_{i+1}) - 2y(x_i) + y(x_{i-1})}{h^2} - \frac{h^2}{12}\left[\frac{y^{(4)}(\xi_i^+) + y^{(4)}(\xi_i^-)}{2}\right]$$

(1*)

Intermediate value theorem implies that there exists some $\xi_i \in (x_{i-1}, x_{i+1})$ such that

$$\frac{y^{(4)}(\xi_i^+) + y^{(4)}(\xi_i^-)}{2} = y^{(4)}(\xi_i)$$

and hence

$$y''(x_i) = \frac{y(x_{i+1}) - 2y(x_i) + y(x_{i-1})}{h^2} - \frac{h^2}{12} y^{(4)}(\xi_i).$$

This is called the centered-difference formula for $y''(x_i)$ [compare the formula in Numerical differentiation].

Similarly we derive centered-difference formula for $y'(x_i)$ as follows:

$$y(x_{i+1}) = y(x_i) + h\,y'(x_i) + \frac{h^2}{2} y''(x_i) + \frac{h^3}{6} y'''(\eta_i^+)$$

for some $\eta_i^+ \in (x_i, x_{i+1})$ and

$$y(x_{i-1}) = y(x_i) - h\,y'(x_i) + \frac{h^2}{2} y''(x_i) - \frac{h^3}{6} y'''(\eta_i^-)$$

for some $\eta_i^- \in (x_{i-1}, x_i)$. By subtracting these two

$$y(x_{i+1}) - y(x_{i-1}) = 2h\,y'(x_i) + \frac{h^3}{3}\left[\frac{y'''(\eta_i^+) + y'''(\eta_i^-)}{2}\right].$$

Using intermediate value theorem

$$\frac{y'''(\eta_i^+) + y'''(\eta_i^-)}{2} = y'''(\eta_i), \quad \text{for some } \eta_i \in (x_{i-1}, x_{i+1})$$

Hence

$$y'(x_i) = \frac{y(x_{i+1}) - y(x_{i-1})}{2h} - \frac{h^3}{6} y'''(\eta_i), \quad \eta_i \in (x_{i-1}, x_{i+1}).$$

1**

Substituting Taylor approximation for $y''(x_i)$ and (3)
$y'(x_i)$ in the equation (1), we get

$$\frac{y(x_{i+1}) - 2y(x_i) + y(x_{i-1})}{h^2} = p(x_i)\left[\frac{y(x_{i+1}) - y(x_{i-1})}{2h}\right] + q(x_i)\,y(x_i) + r(x_i)$$
$$- \frac{h^2}{12}\left[2p(x_i)\,y'''(\eta_i) - y^{(4)}(\xi_i)\right]. \quad \text{—(2)}$$

A finite difference method with truncation error $O(h^2)$ results by using this equation (2) together with boundary conditions $y(a) = \alpha$ and $y(b) = \beta$ as a system of linear equations:

$$w_0 = \alpha, \quad w_{N+1} = \beta.$$

and

$$\left(\frac{-w_{i+1} + 2w_i - w_{i-1}}{h^2}\right) + p(x_i)\left(\frac{w_{i+1} - w_{i-1}}{2h}\right) + q(x_i)\,w_i = -r(x_i) \quad \text{—(3)}$$

for each $i = 1, 2, \ldots N$.

we can rewrite (3) as

$$-\left(1 + \frac{h}{2}p(x_i)\right)w_{i-1} + \left(2 + h^2 q(x_i)\right)w_i - \left(1 - \frac{h}{2}p(x_i)\right)w_{i+1} = -h^2 r(x_i)$$

which can be written as matrix system

$$Aw = b, \text{ where}$$

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_{N-1} \\ w_N \end{bmatrix}, \quad b = \begin{bmatrix} -h^2 r(x_1) + \left(1 + \frac{h}{2}p(x_1)\right)w_0 \\ -h^2 r(x_2) \\ \vdots \\ -h^2 r(x_{N-1}) \\ -h^2 r(x_N) + \left(1 - \frac{h}{2}p(x_N)\right)w_{N+1} \end{bmatrix}$$

and

$$A = \begin{bmatrix} 2+h^2 q(x_1) & -1+\frac{h}{2} p(x_1) & 0 & \cdots & \cdots & 0 \\ -1-\frac{h}{2} p(x_2) & 2+h^2 q(x_2) & -1+\frac{h}{2} p(x_2) & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \\ \vdots & & & \ddots & \ddots & -1+\frac{h}{2} p(x_{N-1}) \\ 0 & \cdots & & & -1-\frac{h}{2} p(x_N) & 2+h^2 q(x_N) \end{bmatrix}$$

The matrix A is tridiagonal $N \times N$ matrix with $A = [a_{ij}]_{1 \le i, j \le N}$.

$$a_{i,i} = 2 + h^2 q(x_i), \quad i = 1, 2, \ldots N.$$

$$a_{i,i+1} = -1 + \frac{h}{2} p(x_i), \quad i = 1, 2, \ldots N-1.$$

$$a_{i+1,i} = -1 - \frac{h}{2} p(x_{i+1}), \quad i = 1, 2, \ldots N-1.$$

Finite difference method for Nonlinear Problems:

Consider

$$y'' = f(x, y, y'), \quad a \le x \le b, \quad y(a) = \alpha, \quad y(b) = \beta.$$

As in the linear case, consider mesh points

$$x_i = a + ih, \quad i = 0, 1, \dots N+1, \qquad h = \frac{b-a}{N+1}.$$

Writing the equation at each interior mesh point x_i, $i = 1, 2, \dots N$.

$$y''(x_i) = f(x_i, y(x_i), y'(x_i)).$$

and using the approximation for $y''(x_i)$ and $y'(x_i)$, we find

$$\frac{y(x_{i+1}) - 2y(x_i) + y(x_{i-1})}{h^2} = f\left(x_i, y(x_i), \frac{y(x_{i+1}) - y(x_{i-1})}{2h} - \frac{h^2}{6} y'''(\eta_i)\right) + \frac{h^2}{12} y^{(4)}(\xi_i)$$

for some ξ_i and η_i in (x_{i-1}, x_{i+1}).

The difference method is obtained by deleting the error terms and using the boundary conditions:

$$w_0 = \alpha, \quad w_{N+1} = \beta.$$

and

$$-\frac{w_{i+1} - 2w_i + w_{i-1}}{h^2} + f\left(x_i, w_i, \frac{w_{i+1} - w_{i-1}}{2h}\right) = 0$$

for each $i = 1, 2, \dots N$.

The $N\times N$ nonlinear system has the form (6)

$$\left.\begin{aligned}
&2w_1 - w_2 + h^2 f\left(x_1, w_1, \frac{w_2-\alpha}{2h}\right) - \alpha = 0\\
&-w_1 + 2w_2 - w_3 + h^2 f\left(x_2, w_2, \frac{w_3-w_1}{2h}\right) = 0\\
&\quad\vdots\\
&-w_{N-2} + 2w_{N-1} - w_N + h^2 f\left(x_{N-1}, w_{N-1}, \frac{w_N-w_{N-2}}{2h}\right) = 0\\
&-w_{N-1} + 2w_N + h^2 f\left(x_N, w_N, \frac{\beta-w_{N-1}}{2h}\right) - \beta = 0.
\end{aligned}\right\} \quad (1)$$

These are N nonlinear equations and N unknowns w_i, $i=1,2\ldots N$ to be found from the equations.

We use Newton's method to solve (1).

A sequence $\{(w_1^{(k)}, w_2^{(k)}\ldots w_N^{(k)})\}$ is generated that converges to the solution of (1) starting from an initial approximation $\{(w_1^{(0)}, w_2^{(0)}, \ldots w_N^{(0)})\}$.

The Newton method for the equations

$$F_j(w_1, w_2\ldots w_N) = 0, \quad j=1,2\ldots N$$

is written as

$$(JF)\left[(w_1^{(k)}, w_2^{(k)}, \ldots w_N^{(k)})^T - (w_1^{(k-1)}, w_2^{(k-1)}, \ldots w_N^{(k-1)})^T\right] =$$

$$= -F(w_1^{(k-1)}, w_2^{(k-1)}, \ldots w_N^{(k-1)}),$$

$k\geq 1$.

$F = (F_1, F_2 \ldots F_N)^T$ and JF is the jacobian of F.

In the case of the system (1) in page (6).
the Jacobian $J(w_1, w_2 \dots w_N)$ is tridiagonal
with ij-th entry

$$J(w_1, w_2 \dots w_N)_{ij} = \begin{cases} -1 + \frac{h}{2} f_{y'}\left(x_i, w_i, \frac{w_{i+1} - w_{i-1}}{2h}\right); & \text{for } i = j-1,\ j = 2, \dots N. \\[2ex] 2 + h^2 f_y\left(x_i, w_i, \frac{w_{i+1} - w_{i-1}}{2h}\right), & \text{for } i = j,\ j = 1, 2, \dots N. \\[2ex] -1 - \frac{h}{2} f_{y'}\left(x_i, w_i, \frac{w_{i+1} - w_{i-1}}{2h}\right), & \text{for } i = j+1,\ j = 1, \dots N-1. \end{cases}$$

where $w_0 = \alpha$ and $w_{N+1} = \beta$.

In the case of linear boundary value problem and in the case of non linear boundary value problem employing newton's method, the finite difference method results in solving a matrix system of the form $Aw = b$, where A is an $N \times N$ tridiagonal matrix. We need good algorithm to solve the matrix system $Aw = b$.

Gaussian Elimination with Backward substitution.

To solve $n\times n$ linear system

$$E_1: a_{11}x_1+a_{12}x_2+ \cdots + a_{1n}x_n = a_{1,n+1}$$
$$E_2 = a_{21}x_1+a_{22}x_2+ \cdots + a_{2n}x_n = a_{2,n+1}$$
$$\vdots$$
$$E_n = a_{n1}x_1+a_{n2}x_2+ \cdots + a_{nn}x_n = a_{n,n+1}.$$

INPUT: number of unknowns and equations n: augmented matrix $A=[a_{ij}]$, $1\le i\le n$ and $1\le j\le n+1$.

OUTPUT: solution $x_1, x_2, \cdots x_n$ or message that the system has no unique solution.

step 1: for $i=1,2,\cdots n-1$, do steps 2-4. (Elimination process)

step 2: let p be ~~smallest the~~ smallest integer with $1\le p\le n$ and $a_{pi}\ne 0$

If no integer p can be found

then OUTPUT ('no unique solution exists'); STOP.

step 3: If $p\ne i$ then perform $(E_p)\leftrightarrow(E_i)$

step 4: for $j=i+1,\cdots n$ do steps 5 and 6.

step 5: set $m_{ji}=a_{ji}/a_{ii}$

step 6: perform $(E_j - m_{ji}E_i)\rightarrow(E_j)$;

step 7: If $a_{nn}=0$, then OUTPUT ('no unique solution')

STOP.

step 8: set $x_n = a_{n,n+1}/a_{nn}$ (start backward substitution)

step 9: for $i=n-1, n-2, \cdots 1$,

set $x_i = \left(a_{i,n+1} - \sum_{j=i+1}^{n} a_{ij}x_j\right)/a_{ii}$

step 10: OUTPUT $(x_1, x_2, \cdots x_n)$;

(Procedure completed successfully)

STOP.

Consider equations:

$E_1: \; a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = a_{1,n+1}$

$E_2: \; a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = a_{2,n+1}$

$\vdots$

$E_n: \; a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = a_{n,n+1}$,

The matrix system corresponding to the above equations is

$$Ax = b,$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}; \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}; \quad b = \begin{bmatrix} a_{1,n+1} \\ a_{2,n+1} \\ \vdots \\ a_{n,n+1} \end{bmatrix}.$$

Define the Augmented matrix $[A; b]$ by

$$[A; b] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & ; & a_{1,n+1} \\ a_{21} & a_{22} & \cdots & a_{2n} & ; & a_{2,n+1} \\ \vdots & & & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} & ; & a_{n,n+1} \end{bmatrix}$$

Gauss-Jordan elimination method consists of three basic types of operations on the equations of a linear system:

(1) Interchanging two equations : $(E_j) \leftrightarrow (E_i)$

(2) Multiplying all the terms of an equation by a nonzero scalar : $(\lambda E_i) \rightarrow (E_i)$; $\lambda \neq 0$.

(3) Adding to one equation a scalar multiple of another : $(E_i + \lambda E_j) \rightarrow E_i$.

Step 5, 6: involve

for $i = 1, 2, \dots n-1$

for $j = i+1, \dots n$,

$$m_{ji} = \frac{a_{ji}}{a_{ii}} \quad \text{--- } 1$$

$$E_j - m_{ji} E_i \rightarrow E_j \quad \text{--- } 2$$

For each i 1 involves $(n-i)$ multiplication/division

$(\because j = i+1, \dots n)$

$\underbrace{1, 2, \dots i}_{i}, \underbrace{i+1, \dots n}_{(n-i)}$

for 2: multiplications

$i=1$; $j=2$, $j=3$, ... $j=n$

$m_{21}E_1$, $m_{31}E_1$, ... $m_{n1}E_1$ → total $(n-1)(n+1)$

multiplications $(n+1)$ $(n+1)$ $(n+1)$

$i=2$; $j=3$, $j=4$, ... $j=n$

$m_{32}E_2$, $m_{42}E_2$ ... $m_{n2}E_2$ → total $(n-2)n$.

n n n

[In E_2, a_{21} entry made zero]

$i=3$; $j=4$, $j=5$, ... $j=n$

$m_{43}E_3$, $m_{53}E_3$ ... $m_{n3}E_3$

$n-1$ $n-1$ $n-1$

[In E_3, a_{31}, a_{32} were made zeros]

total $(n-3)(n-1)$

for general i, we have multiplication in 2,

$$(n-i)(n+2-i)$$

Total multiplication in 1 & 2

$$(n-i)(n+2-i) + (n-i) = (n-i)(n+3-i)$$

2

For (2): addition / subtraction:

$i=1$; $j=2$, $j=3$, - - $j=n$

$E_2 - m_{21}E_1$, $E_3 - m_{31}E_1$ - - $E_n - m_{n1}E_1$ $\longrightarrow$ total

$(n+1)$ $(n+1)$ $(n+1)$ $(n-1)(n+1)$

$i=2$; $j=3$, $j=4$, - - $j=n$ In E_2:

$E_3 - m_{32}E_2$ $E_4 - m_{42}E_2$ - - $E_n - m_{n2}E_2$ $a_{21}=0$;

n n n $\hookrightarrow$ total $(n-2)n$.

for general i, we have number of subtraction in (2)

$$(n-i)(n+2-i)$$

Total addition / subtraction & multiplication / division in (1) & (2) are

$$\sum_{i=1}^{n-1}(n-i)(n+3-i) + \sum_{i=1}^{n-1}(n-i)(n+2-i)$$

$$= \sum_{i=1}^{n-1}\left(n^2-2ni+i^2+3(n-i)\right) + \sum_{i=1}^{n-1}\left(n^2-2ni+i^2+2(n-i)\right)$$

$$= 2\sum_{i=1}^{n-1}(n-i)^2 + 5\sum_{i=1}^{n-1}(n-i)$$

$$= 2\sum_{i=1}^{n-1} i^2 + 5\sum_{i=1}^{n-1} i$$

$$= 2\,\frac{(n-1)n(2n-1)}{6} + 5\cdot\frac{(n-1)n}{2}$$

$$= n(n-1)\left[\frac{4n-2}{6} + \frac{5}{2}\right] = n(n-1)\left[\frac{4n-2+15}{6}\right]$$

$$= \frac{n^2-n}{6}(4n+13) = O(n^3).$$

In the Gauss elimination method, we use diagonal entry, (step 5) ~~divide the~~ a_{ii} of the row E_i, to divide entire E_i, then multiply E_i by $\frac{a_{ji}}{a_{ii}}$ to get $\frac{a_{ji}}{a_{ii}} E_i$, and then subtract it from E_j to make the entries in ith column of each E_j to zero ($j = i+1, \ldots n$). At any of the intermediate step, the diagonal enty $\tilde{a}_{ii}$ in $\tilde{E}_i$ (after doing some steps 5, 6). can become zero. In that case we interchange the row $\tilde{E}_i$ with another row $\tilde{E}_p$ $(p>i)$ where the entry $a_{pi} \neq 0$.

This interchange the rows $(E_p) \leftrightarrow (E_i)$ is called Pivoting strategy:

Consider Examples:

$$\left|\begin{array}{ccc:c} 1 & -1 & 2 & 1 \\ 2 & -2 & 1 & 2 \\ 1 & 2 & 1 & 3 \end{array}\right| \quad \begin{array}{l} E_2 \leftarrow E_2 - 2E_1 \\ E_3 \leftarrow E_3 - E_1 \end{array} \Rightarrow$$

$$\Rightarrow \left|\begin{array}{ccc:c} 1 & -1 & 2 & 1 \\ 0 & 0 & -3 & 0 \\ 0 & 3 & -1 & 2 \end{array}\right| \quad \text{Here } a_{22} = 0; \text{ Pivoting strategy we do } (E_2) \leftrightarrow (E_3)$$

$$\Rightarrow \left|\begin{array}{ccc:c} 1 & -1 & 2 & 1 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & -3 & 0 \end{array}\right| \Rightarrow \begin{array}{l} -3x_3 = 0 \\ 3x_2 - x_3 = 2 \\ x_1 - x_2 + 2x_3 = 1 \end{array}$$

$$\Rightarrow x_3 = 0; \quad x_2 = 2/3; \quad x_1 = 1 + x_2 = 1 + 2/3 = 5/3$$

Iteration Methods:

The Gauss - Jacobi method: (simultaneous displacements):

Objective is to find solution x of the system $Ax=b$.

we rewrite the system $Ax=b$. by

① $$x_i = \frac{1}{a_{ii}}\left\{ b_i - \sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j \right\}, \quad i=1,2,\ldots n.$$

assuming all $a_{ii}\neq 0$. Define the iteration by

② $$x_i^{(m+1)} = \frac{1}{a_{ii}}\left\{ b_i - \sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j^{(m)} \right\}, \quad i=1,2,\ldots n, \quad m\geq 0$$

and assume that the initial guess $x_i^{(0)}$, $i=1,\ldots n$ are given.

To analyse the convergence, let $e^{(m)} = x - x^{(m)}$, $m\geq 0$.

from ① & ②.

$$e_i^{(m+1)} = -\sum_{\substack{j=1\\ j\neq i}}^{n} \frac{a_{ij}}{a_{ii}}\, e_j^{(m)}, \quad i=1,2,\ldots n, \quad m\geq 0.$$

For a vector $y=(y_1, y_2, \ldots y_n)$, define

$$\|y\|_\infty = \max_{1\leq i\leq n} |y_i|$$

Then

$$|e_i^{(m+1)}| \le \sum_{\substack{j=1\\ j\neq i}}^{n} \left|\frac{a_{ij}}{a_{ii}}\right| \|e^{(m)}\|_\infty$$

Define

$$\mu = \max_{1\le i\le n} \sum_{\substack{j=1\\ j\neq i}}^{n} \frac{|a_{ij}|}{|a_{ii}|}$$

Then

$$|e_i^{(m+1)}| \le \mu \|e^{(m)}\|_\infty$$

Since the right hand side is independent of i,

$$\|e^{(m+1)}\|_\infty \le \mu \|e^{(m)}\|_\infty$$

If $\mu < 1$, then $e^{(m)} \to 0$ as $m \to \infty$.

whith a linear rate bounded by μ, and

$$\|e^{(m)}\|_\infty \le \mu^m \|e^{(0)}\|_\infty.$$

In order for $\mu < 1$ to be true, the matrix A must be diagonally dominant, that is

$$\sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| < |a_{ii}|, \quad i = 1, 2 .. n.$$

example: The matrix $A = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 3 \end{bmatrix}$

is diagonally dominant.