

Contents

I	MATH 572	2
1	Course Notes	3
1.1	01/01/16	4
1.1.1	Floating Point Numbers	4
1.1.2	Finite Difference Approximation:	4
1.1.3	Truncation Errors:	5
1.2	01/12/16	6
1.2.1	Method of Undetermined coefficients	6
1.2.2	Higher Derivatives	7
1.2.3	Solving Boundary Value Problems (B.V.P.'s)	8
1.3	01/14/16	9
1.3.1	Continuation from 01/12/16	9
1.3.2	Local Truncation Error (L.T.E)	9
1.3.3	Global Error	10
1.3.4	Deferred Correction Method:	11
1.3.5	Stability:	11
1.3.6	Consistency:	11
1.3.7	Convergence:	11
1.4	01/19/16	12
1.4.1	Continuation from 01/14/2016	12
1.4.2	More General P.D.E.'s	12
1.4.3	More General Boundary Value Problems	13
1.4.4	Extrapolation Method:	14
2	Homework	15
2.1	Homework #1	15

Part I
MATH 572

1 Course Notes

1.1 01/01/16

What is an algorithm? - sequence of steps to accomplish a given task. Numerical algorithm: differentiating, integration, solving differential equations $f = u''$. Finite difference methods (FDM): They are straight forward to implement, but they suffer from several drawbacks. This course is mainly about FD methods and the analysis.

1.1.1 Floating Point Numbers

There are a finite amount of numbers

$$x = \pm(0.d_1d_2\dots d_n)\beta^{\pm e}, \quad (1)$$

where $(0.d_1d_2\dots d_n)$ is the mantissa, β is the base, and e is the exponent. $d_1 \neq 0$ by convention. Example of how bits are given in a number are shown in Table 1

Machine	Sign	Exponential sign	Mantissa	Exponent
32-bit	\pm	\pm	n=23	m=7
64-bit	\pm	\pm	n=52	m=10

Table 1: Floating point numbers for different machines

Example: Consider a computer with $\beta = 2$, $n = 4$, and $m = 3$. What is x_{max} =?

$$\begin{aligned} x_{max} &= (0.1111) \cdot 2^3 \\ &= [1 \times 2^{-1} + 1 \times 2^{-2} + 1 \times 2^{-3} + 1 \times 2^{-4}] \cdot 2^3 \\ &= 4 + 2 + 1 + \frac{1}{2} = 7.5, \end{aligned} \quad (2)$$

and x_{min} =?

$$\begin{aligned} x_{min} &= (0.1000) \cdot 2^{-3} \\ &= 2^{-4} = \frac{1}{32} = 0.0625. \end{aligned} \quad (3)$$

Shaded area is where we can represent numbers, but not all the numbers in the shaded area can be represented. ***Add in figure to show this*** Roundoff Error:

$$|x - \text{fl}(x)|. \quad (4)$$

This is from loss of significant digits, and techniques to overcome, etc...

1.1.2 Finite Difference Approximation:

Given $u(x)$, find $u'(x)$

$$u(x) = \sin(x) \rightarrow u'(x) = \cos(x). \quad (5)$$

What if u is complicated or what if $u(x)$ is not given analytically? ***insert pictures***

$$D_+ u(x) = \frac{u(x+h) - u(x)}{h} \quad (6)$$

for some small parameter. Similarly we can define

$$D_-u(x) = \frac{u(x) - u(x-h)}{h}. \quad (7)$$

Both D_+u and D_-u are first order accurate approximations to $u'(x)$, that is

$$\lim_{h \rightarrow 0} |D_+u(x) - u'(x)| = \mathcal{O}(h), \quad (8)$$

if h goes down by $\frac{1}{2}$ then error goes down by $\frac{1}{2}$ as well. Define

$$D_0u = \frac{D_+u + D_-u}{2} = \frac{u(x+h) - u(x-h)}{2h}. \quad (9)$$

$$\lim_{h \rightarrow 0} |D_0u(x) - u'(x)| = \mathcal{O}(h^2), \quad (10)$$

where the error reduction in this case is by a factor of h^2 .

1.1.3 Truncation Errors:

Take a Taylor series expansion of $u(x+h)$

$$u(x+h) = u(x) + hu'(x) + \frac{h^2}{2!}u''(x) + \dots \quad (11)$$

and of $u(x-h)$

$$u(x-h) = u(x) - hu'(x) + \frac{h^2}{2!}u''(x) + \dots \quad (12)$$

From Eqn. 11 we get:

$$D_+u(x) = \frac{u(x+h) - u(x)}{h} = u'(x) + \frac{h}{2!}u''(x) + \dots \quad (13)$$

$$u'(x) = D_+u(x) - \frac{h}{2!}u''(x) + \dots \quad (14)$$

$$u'(x) = D_+u(x) + \mathcal{O}(h) \quad (15)$$

and from Eqn. 12 we get:

$$D_-u(x) = \frac{u(x) - u(x-h)}{h} = u'(x) - \frac{h}{2!}u''(x) + \dots, \quad (16)$$

therefore

$$u'(x) = D_-u(x) + \mathcal{O}(h). \quad (17)$$

Then we have

$$D_0u(x) = \frac{u(x+h) - u(x-h)}{2h} = u'(x) + 2 \cdot \frac{h^3}{3! \cdot 2h}u'''(x) + \dots = u'(x) + \frac{h^2}{3!}u'''(x) + \dots, \quad (18)$$

therefore we have

$$u'(x) = D_0u(x) + \mathcal{O}(h^2), \quad (19)$$

which is second order accurate. This is also known as a center difference scheme. Deriving finite difference formulas.

Example: Given $u(x)$, $u(x-h)$, and $u(x+h)$, let

$$D_2u(x) = au(x) + bu(x-h) + cu(x+h), \quad (20)$$

find a, b, c that gives best possible approximations to $u'(x)$. Subsitute the Taylor series expansion for $u(x-h)$, $u(x-2h)$ about the point x :

$$D_2u(x) = (a+b+c)u(x) - (b+2c)hu'(x) + \frac{1}{2}(b+4c)h^2u''(x) - \frac{1}{6}(b+8c)h^3u'''(x) + \dots \quad (21)$$

Then we enforce that

$$\begin{cases} a+b+c &= 0 \\ -(b+2c)h &= 1, \\ \frac{1}{2}(b+4c)h^2 &= 0 \end{cases} \quad (22)$$

then solve for the coefficients to get

$$\begin{cases} a &= \frac{3}{2}h \\ b &= -\frac{2}{h} \\ c &= \frac{1}{2}h \end{cases} \quad (23)$$

We cannot enforce anything else because we only have three unknowns, therefore

$$D_2u(x) = \frac{1}{2h} [3u(x) - 4u(x-h) + u(x-2h)]. \quad (24)$$

Order of accuracy is ?? Leading order term

$$= -\frac{1}{6}(b+8c)h^3u'''(x) = -\frac{1}{6}\left(-\frac{2}{h} + 8 \cdot \frac{1}{2h}\right)h^3u'''(x) = \frac{1}{3}h^2u'''(x), \quad (25)$$

therefore

$$u'(x) = D_2u(x) + \mathcal{O}(h^2). \quad (26)$$

Function evaluations are always expensive. Use first order method with not the most accurate data set. Need to way both accuracy and stability of algorithm.

1.2 01/12/16

1.2.1 Method of Undetermined coefficients

- Suppose $u(x_1)$, $u(x_2)$, $u(x_3)$ are given
- Find an interpolating polynomial $p(x)$
- Compute $p'(x) \approx u'(x)$

Second order derivative $u''(x)$

$$D^2u = \begin{cases} D_-(D_+u(x)) \\ D_+(D_-u(x)) \\ \hat{D}_0(\hat{D}_0u(x)) \end{cases} \quad (27)$$

1.2.3 Solving Boundary Value Problems (B.V.P.'s)

Starting with

$$u''(x) = f(x); u(0) = \alpha : u(1) = \beta. \quad (35)$$

We first discreteize the domain,

$$x_j = jh, \text{ for } j = 0, 1, \dots, m+1, \quad (36)$$

then write the difference equations and let

$$u_j = u(x_j) = u(jh), \quad (37)$$

given u_0, u_m, f , compute $\{u_1, u_2, \dots, u_m\}$. Since the PDE is valid at all points in the domain, each interior node will give one equation. Then we get m equations and unknowns that can be solved for. Consider x_j , conver the differential equation into a difference equation using F.D. approximation.

$$u'' = f \Big|_{x=x_j} \rightarrow \frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} = f_j, \quad (38)$$

which is valid for all $j = 1, 2, \dots, m$. Consider $j = 1$ the D.E. is

$$f_1 = \frac{u_2 - 2u_1 + u_0}{h^2}, u_0 = \alpha \quad (39)$$

then we get

$$\frac{u_2 - 2u_1}{h^2} = f_1 - \frac{\alpha}{h^2}, \quad (40)$$

and similarly

$$\frac{-2u_m + u_{m+1}}{h^2} = f_m - \frac{\beta}{h^2}. \quad (41)$$

Convert these D.E.'s into a matrix equation

$$\frac{1}{h^2} \begin{bmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & 1 & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & 1 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_m \end{bmatrix} = \begin{bmatrix} f_1 - \frac{\alpha}{h^2} \\ f_2 \\ \vdots \\ \vdots \\ f_m - \frac{\beta}{h^2} \end{bmatrix}, \quad (42)$$

which is in the form of $AU = F$, where A is a tridiagonal matrix, and this is formed with Dirichlet Boundary conditions. We need to look at

- local truncation error
- Global truncation error $|u(x) - U(x)|$, where $u(x)$ is the exact solution and $U(x)$ is the computed solution.
- Stability
- Consistency
- Convergence.

1.3 01/14/16

1.3.1 Continuation from 01/12/16

Neumann's B.C.'s are

$$u''(x) = f(x); u'(0) = \alpha; u'(1) = \beta. \quad (43)$$

Then discretize the domain

$$0, h, 2h, (m+1)h \quad (44)$$

with the Difference equations: at $x_j = jh$

$$\frac{1}{h^2} [U_{j+1} - 2U_j + U_{j-1}] = f_j \quad (45)$$

and at $x_0 = 0$

$$u'(0) = \alpha \quad (46)$$

Use a first order Finite Difference approximation on u'

$$\frac{u_1 - u_0}{h} = \alpha \quad (47)$$

or use a second order approximation for u' (or apply polynomial approximation, or method of undetermined coefficients on u_0, u_1, u_2)

Similarly at x_{m+1}

$$\frac{u_{m+1} - u_m}{h} = \beta \quad (48)$$

or apply the second order approx. The matrix equation is

$$\frac{1}{h^2} \begin{bmatrix} -h & h & 0 & \cdots & \cdots & \cdots & 0 \\ 1 & -2 & 1 & & & & \vdots \\ 0 & 1 & -2 & 1 & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & & 1 & -2 & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & -h & h \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ \vdots \\ \vdots \\ u_m \\ u_{m+1} \end{bmatrix} = \begin{bmatrix} \alpha \\ f_1 \\ \vdots \\ \vdots \\ \vdots \\ f_m \\ \beta \end{bmatrix} \quad (49)$$

$$AU = f. \quad (50)$$

A matrix is not invertable in this form (it had a null space), therefore this is a problem.

1.3.2 Local Truncation Error (L.T.E)

Loosely speaking, error due to truncating Talyor series, $u'' = f$, with τ being the notatoin for the L.T.E. In vetor form

$$\tau = AU \hat{=} f \quad (51)$$

at x_j

$$\tau = \frac{1}{h^2} [u(x_{j-1}) - 2u(x_j) + u(x_{j+1})] - f(x_j) \quad (52)$$

$$\tau = \left[u''(x_j) + \frac{1}{12} h^2 u''''(x_j) + \mathcal{O}(h^4) \right] - f(x_j). \quad (53)$$

By expanding $u(x_{j-1})$ and $u(x_{j+1})$ via Taylor series about x_j therefore

$$\tau_j = \frac{1}{12} h^2 u''''(x_j) + \mathcal{O}(h^4) \implies \tau_j = \mathcal{O}(h^4). \quad (54)$$

1.3.3 Global Error

The global error is defined by

$$E = U - \hat{U}, \quad (55)$$

where E is the global error, U is the numerical approximate solution and \hat{U} is the exact solution

$$\hat{U} = \begin{bmatrix} u(x_1) \\ u(x_2) \\ \vdots \\ u(x_m) \end{bmatrix}. \quad (56)$$

Recall from before that $\tau = A\hat{U} - f$

Question: If we know τ exactly, can we obtain E ?

$$AE = AU - A\hat{U}$$

$$AE = f - A\hat{U}$$

$$AE = -\tau, \quad (57)$$

which can be used for multigrid and deferred correction methods. If A is invertible then we can find E . Deferred correction methods estimate τ to find the error, then use the found error to improve the solution.

$$AE = -\tau$$

can be integrated as

$$u'' = f, \quad (58)$$

with $u(0) = \alpha$ and $u(1) = \beta$. Then we have

$$e'' = -\tau, \quad (59)$$

such that $e(0) = e(1) = 0$. For this Dirichlet Problem

$$\tau_j = \frac{1}{12} h^2 u''''(x_j) + \mathcal{O}(h^4) \quad (60)$$

$$\implies \tau_j = \frac{1}{12} h^2 f''(x_j) + \mathcal{O}(h^4). \quad (61)$$

Therefore we know approximatley what τ_j is.

1.3.4 Deferred Correction Method:

1. Solve $AU = f \rightarrow \mathcal{O}(h^2)$ scheme
2. Solve $AE = \frac{1}{12}h^2 f''$
3. Set $U \rightarrow U + E$

with $\rightarrow A$ on the $\mathcal{O}(h^2)$ scheme. The price paid is solving $AE = ()$, can we avoid this? For this example, yes! Add 1) and 2) to get:

$$A(U + E) = f + \frac{1}{12}h^2 f'', \quad (62)$$

but in general this can not be avoided.

1.3.5 Stability:

$$A^h E^h = -\tau^h \quad (63)$$

and we want $E^h \rightarrow 0$ as $h \rightarrow 0$.

Definition: we say that our F.D. method is stable if $(A^h)^{-1}$ exists for all h sufficiently small ($h < h_0$) and if there exists a constant c such that

$$\|(A^h)^{-1}\| \leq c \quad \forall h < h_0 \quad (64)$$

which is defined for by any norm type.

1.3.6 Consistency:

Consistency is for error truncation. Our F.D. method is consistent with the differential equation and boundary conditions if $\|\tau^h\| \rightarrow 0$ as $h \rightarrow 0$. Typically

$$\|\tau^h\| \sim \mathcal{O}(h^p) \quad (65)$$

where p is the power. As long as truncation error is something like this then consistency of our method is met.

1.3.7 Convergence:

Convergence is important in regards to global error. Our F.D. method is convergent if $\|E^h\| \rightarrow 0$ as $h \rightarrow 0$. The Fundamental theorem of Finite Difference Method is

$$\text{Consistency} + \text{Stability} = \text{Convergence}. \quad (66)$$

If we have stability and consistency, then our method will have convergence. Proof:

$$A^h E^h = -\tau^h \quad (67)$$

$$E^h = -(A^h)^{-1} \tau^h, \quad (68)$$

since $(A^h)^{-1}$ exists

$$\|E^h\| = \|(A^h)^{-1} \tau^h\| \quad (69)$$

$$\|E^h\| \leq \|(A^h)^{-1}\| \cdot \|\tau^h\| \quad (70)$$

and from the stability condition in Eqn. 64 we get

$$\|E^h\| \leq c \cdot \|\tau^h\| \quad (71)$$

therefore

$$\|E^h\| \rightarrow 0 \text{ as } h \rightarrow 0 \quad (72)$$

because $\|\tau^h\| \rightarrow 0$ as $h \rightarrow 0$ due to the consistency condition. More importantly,

$$\mathcal{O}(h^p) \text{ L.T.E.} + \text{Stability} \implies \mathcal{O}(h^p) \text{ for global error.}$$

1.4 01/19/16

1.4.1 Continuation from 01/14/2016

$$\begin{bmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & 1 & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & 1 & 2 \end{bmatrix} \text{ is a symmetric and negative definite matrix.}$$

1.4.2 More General P.D.E.'s

Starting with

$$f(x) = (k(x)u'(x))'; \quad u(0) = \alpha; \quad u(1) = \beta \quad (73)$$

From this we get that

$$f = ku'' + k'u' \quad (74)$$

which is similar to our previous strategy and at x_i

$$u''(x_i) = \frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} \quad (75)$$

$$u'(x_i) = \frac{u_{i+1} - u_{i-1}}{2h} \quad (76)$$

At which x_1 we get the following difference equation, and we imply k is a given quantity, therefore

$$f_i = k_i \left(\frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} \right) + k_i' \left(\frac{u_{i+1} - u_{i-1}}{2h} \right), \quad (77)$$

where $k_i = k(x_i)$, and $f_i = f(x_i)$. From this we can set up the matrix equation, $AU = F$, where

$$A = \frac{1}{h^2} \begin{bmatrix} -2k_1 & k_1 + h \frac{k_1'}{2} & & & \\ k_2 - h \frac{k_2'}{2} & -2k_2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots \\ & & & \ddots & -2k_n \end{bmatrix}, \quad (78)$$

which is not a symmetric matrix. If we consider halfway points inbetween $i-1, i$, and $i+1$ we can have the following

$$f = (ku')', \quad (79)$$

where

$$k_{i+1/2}u'(x_{i+1/2}) = k_{i+1/2} \left(\frac{u_{i+1} - u_i}{h} \right), \quad (80)$$

and

$$k_{i-1/2}u'(x_{i-1/2}) = k_{i-1/2} \left(\frac{u_i - u_{i-1}}{h} \right). \quad (81)$$

Now we can apply the difference scheme at x_i to get

$$f(x_i) = \frac{k_{i+1/2}u'(x_{i+1/2}) - k_{i-1/2}u'(x_{i-1/2})}{h}, \quad (82)$$

then we can get

$$f_i = \frac{1}{h^2} [k_{i+1/2}(u_{i+1} - u_i) - k_{i-1/2}(u_i - u_{i-1})], \quad (83)$$

$$f_i = \frac{1}{h^2} [k_{i+1/2}u_{i+1} - (k_{i+1/2} + k_{i-1/2})u_i + k_{i-1/2}u_{i-1}]. \quad (84)$$

This then gives us the following A matrix

$$A = \frac{1}{h^2} \begin{bmatrix} -(k_{1+1/2} + k_{1-1/2}) & k_{1+1/2} & & & & \\ k_{2-1/2} & -(k_{2+1/2} + k_{2-1/2}) & k_{2+1/2} & & & \\ & k_{3-1/2} & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & k_{n-1+1/2} \\ & & & & k_{n-1/2} & -(k_{n+1/2} + k_{n-1/2}) \end{bmatrix}, \quad (85)$$

which is a symmetric and negative definite matrix, all eigen values are negative.

1.4.3 More General Boundary Value Problems

Similarly F.D.M. can be applied to more general boundary value problems (B.V.P.'s), such as:

$$f(x) = a(x)u''(x) + b(x)u'(x) + c(x)u(x). \quad (86)$$

Higher order method

$$f = u''; u(0) = \alpha; u(1) = \beta. \quad (87)$$

Simply apply as higher order approximation to all the derivatives involved. Here, $u''(x)$ is a 4th order approximation

$$\frac{1}{12h^2} [-u_{i-2} + 16u_{i-1} - 30u_i + 16u_{i+1} - u_{i+2}] \quad (88)$$

1.4.4 Extrapolation Method:

We have two grids, one course grid and one fine grid, where u_j are points on the course grid and v_j are points on the fine grid. Starting from $f = u$; $u(0) = \alpha$; $u(1) = \beta$, we know that u_j solved via central difference schemes satisfies the estimate:

$$\text{for } j = 1, 2, \dots, m \quad (89)$$

$$u_j - u(jh) = c_2 h^2 + c_4 h^4 + c_6 h^6 + \dots = \text{Error Equation}, \quad (90)$$

and

$$v_{2j} - u(jh) = c_2 \left(\frac{h}{2}\right)^2 + c_4 \left(\frac{h}{2}\right)^4 + c_6 \left(\frac{h}{2}\right)^6 + \dots, \quad (91)$$

which the error is reduced by a factor of $\frac{1}{4}$ for the fine grid. Multiply Eqn. 91 by 4 and we get

$$4v_{2j} - 4u(jh) = c_2 h^2 + 4c_4 \left(\frac{h}{2}\right)^4 + 4c_6 \left(\frac{h}{2}\right)^6 + \dots \quad (92)$$

Now subtract Eqn. 92 from Eqn. 90 and we get

$$u_j - 4v_{2j} + 3u(jh) = c_4 \left(1 - \frac{1}{4}\right) h^4 + \mathcal{O}(h^6), \quad (93)$$

then divide by -3

$$\frac{4v_{2j} - u_j}{3} - u(jh) = -\frac{1}{3} c_4 \frac{3}{4} h^4 + \mathcal{O}(h^6), \quad (94)$$

where $\bar{u}_j = \frac{4v_{2j} - u_j}{3}$, therefore \bar{u}_j is a $\mathcal{O}(h^4)$ accurate solution, so for the Extrapolation algorithm

- solve for u_j on the course grid using central difference scheme
- solve for v_j on the fine grid using central difference scheme
- set $\bar{u}_j = \frac{4v_{2j} - u_j}{3}$

We can also use nonuniform course and fine grids. What happens if all the data points are used for each difference equation? Use the spectral method. Higher the order, steeper the slope.

$$f(x) = \sum_{j=0}^n c_j x^j \quad (95)$$

is an unstable approximation. In general, approximately data given on a regular grid with monic polynomial represented leads to inaccurate approximations \rightarrow Runge phenomenon??

2 Homework

2.1 Homework #1