CHAPTER 2

Steady States and BVP

Consider the IBVP

$$\begin{cases} u_t - u_{xx} = x \\ u(0,t) = 2 \\ u(1,t) = 3 \\ u(x,0) = g(x) \end{cases}$$

Soln: $$\boxed{u(x,t) = \sum_{n=1}^{\infty} a_n \sin(n\pi x)\, e^{-(n\pi)^2 t} + \frac{x^3}{6} + \frac{5}{6}x + 2}$$

The coefficients a_n are obtained from the Fourier sine Series of $g(x) - \frac{x^3}{6} + \frac{5}{6}x + 2$

Obviously, when $t \to \infty$, $u(x,t) \to \frac{x^3}{6} + \frac{5}{6}x + 2$

Which is the soln. for the equillibrium problem

$$\begin{cases} \hat{u}''(x) = x \\ u(0) = 2 \\ u(1) = 3 \end{cases}$$

2.4 Simple Finite Difference Method.

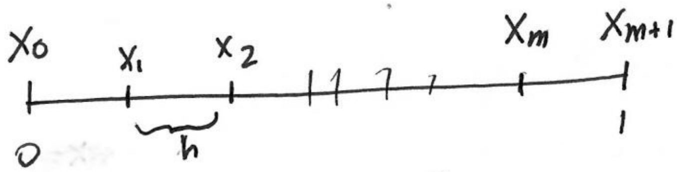
Consider

$$\begin{cases} u''(x) = f(x), & 0 < x < 1 \quad (1) \\ u(0) = \alpha, \quad u(1) = \beta & \quad (2) \end{cases}$$

If $f(x)$ is twice integrable, it can be easily solve by direct integration.

An alternative is to find a numerical solution.

Step 1: Discretization of domain



$m+2$ points
$m+1$ Subintervals
Uniform grid
step size $= h$.

Let's use $\boxed{U_j \simeq u(x_j)}$

and $U_0 = u(0) = \alpha$, $\quad U_{m+1} = u(x_{m+1}) = u(1) = \beta$

Our goal is to find: $U_1, U_2, .., U_m$. "m" unknowns.

The equation (1) can be approximated by the discrete equation:

$$\boxed{D^2 U_j = \frac{U_{j+1} - 2U_j + U_{j-1}}{h^2} = f(x_j) = f_j, \quad j = 1, \dots m.} \quad (3)$$

Set of equations for unkws $v_1, .., v_m$.

$j=1$ $$\frac{v_2 - 2v_1 + \overset{=\alpha}{v_0}}{h^2} = f(x_1) \Rightarrow \frac{v_2 - 2v_1}{h^2} = f(x_1) - \frac{\alpha}{h^2}$$

$j=2$ $$\frac{v_3 - 2v_2 + v_1}{h^2} = f(x_2)$$

$\vdots$

$j=i$ $$\frac{v_{i+1} - 2v_i + v_{i-1}}{h^2} = f(x_i)$$

$\vdots$

$j=m$ $$\frac{\overset{=\beta}{v_{m+1}} - 2v_m + v_{m-1}}{h^2} = f(x_m) \Rightarrow \frac{-2v_m + v_{m-1}}{h^2} = f(x_m) - \frac{\beta}{h^2}$$

Therefore, the unkws: $v_1 .., v_m$ can be obtained as the soln. of the following linear system

$$\underbrace{\frac{1}{h^2}\begin{bmatrix} -2 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & -2 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ 0 & 0 & \cdots & & 0 & 1 & -2 \end{bmatrix}}_{\equiv A} \underbrace{\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ \vdots \\ v_m \end{bmatrix}}_{\equiv \vec{v}} = \underbrace{\begin{bmatrix} f(x_1) - \frac{\alpha}{h^2} \\ f(x_2) \\ \vdots \\ f(x_{m-1}) \\ f(x_m) - \frac{\beta}{h^2} \end{bmatrix}}_{\equiv \vec{F}}$$

Tridiag. linear system that has a unique soln. for any f, α, β, and $h>0$

or $\boxed{A\vec{v} = \vec{F}}$

Error in the approximation

Define

$$\hat{U} \equiv \begin{bmatrix} u(x_1) \\ \vdots \\ u(x_m) \end{bmatrix}, \text{ then } \vec{E} \equiv U - \hat{U} = \begin{bmatrix} u(x_1) - U_1 \\ \vdots \\ u(x_m) - U_m \end{bmatrix}$$

Error vector
global error

Goal is to obtain bound for magnitude of E.

We need norms

a) <u>Max norm:</u> or Inf. norm

$$\|E\|_\infty \equiv \max_{1 \le j \le m} |E_j| = \max_{1 \le j \le m} |u(x_j) - U_j|$$

b) <u>1-Norm:</u>

$$\|E\|_1 \equiv h \sum_{j=1}^{m} |E_j|$$

discretization of $\int_{x_1}^{x_{max}} |E(x)| dx =$

$E_j = E(x_j) \equiv U_j - u(x_j)$

c) <u>2-Norm:</u>

$$\|E\|_2 = \left(h \sum |E_j|^2 \right)^{1/2} \leftarrow \left[\int_{x_1}^{x_{max}} E^2(x)\, dx \right]^{1/2}$$

Grid function norms.

Key steps for estimating error in FD approx:

1) Compute LTE.
2) Use stability to bound global error in terms of LTE.

$$\boxed{E = U - \hat{U}}$$

<u>Rmk:</u> LTE: It has to do with the error in FD approx. of derivatives in equation.

Sections 2.4–2.9 (Leveque).

Fundamental Concepts in Numerical Analysis of Numerical Solutions of Differential Equations: Consistency, Convergence, and Stability.

These concepts will be illustrated using the discrete approximation of the model problem in Leveque chapter 2.

$$\rightarrow \begin{cases} u''(x) = f(x), \quad 0 < x < 1 & (1) \\ u(0) = \alpha, \quad u(1) = \beta & (2) \end{cases}$$

$$h \equiv \frac{1}{m+1}$$

uniform partition.

Using centered difference approx. of $O(h^2)$ for $u''(x)$,

(1) Can be approx. by

$$\boxed{D^2 U_j = \underbrace{\frac{U_{j+1} - 2U_j + U_{j-1}}{h^2}} = f(x_j), \quad j = 1, 2, .., m} \quad (3)$$

and (2) implies $\quad U_0 = \alpha, \quad U_{m+1} = \beta \qquad (4)$

Also, $\boxed{U_j \simeq u(x_j).}$

The discrete equations (3)-(4) can be represented as a linear algebraic System of equations:

$$\boxed{A\,U = \vec{F}} \qquad (5)$$

The equ. (5) is also Called a numerical method, where

$$A = \frac{1}{h^2}\begin{bmatrix} -2 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & -2 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & & \\ \vdots & & & & 1 & -2 & 1 \\ 0 & \cdots & & & & 1 & -2 \end{bmatrix}, \quad \vec{U} = \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_m \end{bmatrix}, \quad \vec{F} = \begin{bmatrix} f(x_1) - \alpha/h^2 \\ f(x_2) \\ \vdots \\ f(x_{m-1}) \\ f(x_m) - \beta/h^2 \end{bmatrix}$$

We will also define the vector

$$\hat{U} = \begin{bmatrix} u(x_1) \\ u(x_2) \\ \vdots \\ u(x_m) \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix}$$

Def.- (Global error)

The difference between the two vectors $\vec{E}$

$$\boxed{\vec{E} = \vec{U} - \hat{U} = \begin{bmatrix} U_1 - u_1 \\ U_2 - u_2 \\ \vdots \\ U_m - u_m \end{bmatrix}}$$

is Called the <u>global error</u>.

The purpose of the numerical method (5) is to obtain an approx. of the exact solution $u(x)$ of (1)-(2). This approx. is measured in terms of some norm of the global error $\vec{E}$. For example,

$$\boxed{\|\vec{E}(h)\| = \max_{1\le j\le m} |U_j - u_j|}$$

or

$$\boxed{\|\vec{E}(h)\| = \left[h \sum_{j=1}^{m} |U_j - u_j|^2 \right]^{1/2}}$$

The goal for a numerical method such as (5) is to provide a "good" approx. to $u(x)$. In particular, it is desirable that $\vec{U}$ (numerical approx.) becomes "better" as the "grid is refined" (h is decrease). For this, we introduce a new concept

Def.- (Convergence).

A numerical method such as (5) is said to be convergent if

$$\boxed{\|E(h)\| \longrightarrow 0 \quad \text{as } h \to 0.} \tag{6}$$

Notice that the vector $\vec{E}$ depends on h. Actually, the size of the vector $\vec{E}$ grows when h decreases.

Equation (3) can be written as

$$\frac{v_{j+1} - 2v_j + v_{j-1}}{h^2} - f(x_j) = 0 \qquad (5.1)$$

Replacing the exact solution $u(x)$ of (1)-(2) evaluated at x_j into (5.1) for $j = 1, 2, \ldots, m$

$$\frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} - f(x_j) = u''(x) + \frac{h^2}{12} u^{(4)}(x_j) + O(h^4) - f(x_j)$$

$$= \frac{h^2}{12} u^{(4)}(x_j) + O(h^4) = T_j \qquad (5.2)$$

Since, $\quad u''(x) - f(x_j) = 0.$

T_j is called the local truncation error.

<u>Definition</u>

The Local truncation T_j is the remainder term obtained after substitution of the exact soln. into the corresponding discrete equation.

In our case, $\quad T_j = O(h^2). \qquad (5.3)$

Equation (5.2) is also equivalent to

$$\underbrace{\left[\frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} - f(x_j)\right]}_{\text{discrete equation}} - \underbrace{\left[u''(x_j) - f(x_j)\right]}_{\text{Continuous equation}} = \tau_j$$

So τ_j is also the difference between the discrete equation and the continuous equation using the exact solution at the grid points.

Using the matrix A, we obtain the vector associated to (5.2). In fact,

$$A\hat{U} - \vec{F} = \vec{\tau}$$

where $\vec{\tau} = \begin{bmatrix} \tau_1 \\ \vdots \\ \tau_m \end{bmatrix}$

Since the size of the matrix A and vectors $\vec{U}$, $\hat{U}$, and $\vec{\tau}$ depends on h, we will express the discrete equations as

$$A^h \vec{U}^h = \vec{F}^h \quad \text{and} \quad A^h \hat{U}^h - \vec{F}^h = \vec{\tau}^h$$

Def.- (Consistency)

The numerical method (5) is consistent with the continuous differential equation and boundary conditions if

$$\boxed{\|\vec{\tau}^h\| \xrightarrow[h\to 0]{} 0}$$

Remarks: i) A superscript "h" has been added to indicate the dependence of the local truncation error vector on "h".

ii) From (5.3), it's seen that $A^h\hat{U}^h = \vec{F}^h$ is consistent.

Def.- (Stability)

Consider a numerical method approximating the solution of a linear BVP given by a sequence of matrix equations

$$A^h \vec{U}^h = \vec{F}^h \qquad (6.1)$$

where h is step size. We will say that the numerical method (6.1) is stable if

i) $\boxed{(A^h)^{-1} \text{ exists}}$ for all $h < h_0$

ii) There is a C such that

$$\boxed{\|(A^h)^{-1}\| \le C, \quad \text{for all } h < h_0.}$$

Remark: The symbol $\|(A^h)^{-1}\|$ denotes the norm of a matrix A^h. A review of this concept is provided in the next section 2.10.

An important relationship between Stability, Consistency and Convergence exists. The next theorem exhibits it.

Theorem. (Fundamental theorem of finite difference method)

Consider a numerical method

$$A^h \vec{U}^h - \vec{F}^h = 0 \qquad (8.1)$$

If (8.1) is consistent with a linear BVP and (8.1) is also stable, then it is convergent.

Proof:- We want to prove convergence : $\|E(h)\| \xrightarrow[h\to 0]{} 0$

By definition

$$A^h \vec{U}^h = \vec{F}^h \quad (\text{Equ. (5) in page 2})$$

Also, $$A^h \hat{U}^h = \vec{F}^h + \vec{\tau}^h \quad (\text{Equ. } 5.2)$$

Therefore, $A^h \vec{U}^h - A^h \hat{U}^h = A^h(\vec{U}^h - \hat{U}^h) = -\vec{\tau}^h$

$\Rightarrow A^h \vec{E}^h = -\vec{\tau}^h$

If $(A^h)^{-1}$ exists (stab.)

$\Rightarrow \vec{E}^h = (A^h)^{-1} \vec{\tau}^h$

$\Rightarrow \|\vec{E}^h\| \leq \|(A^h)^{-1}\| \, \|\vec{\tau}^h\| \leq C \|\vec{\tau}^h\|$

Now, consistency implies

$$\|\vec{\tau}^h\| \xrightarrow[h\to 0]{} 0$$

Therefore, $\|\vec{E}^h\| \xrightarrow[h\to 0]{} 0$

but not only that if

$$\|\vec{\tau}^h\| = \mathcal{O}(h^p) \Rightarrow \|\vec{E}^h\| = \mathcal{O}(h^p).$$

In other words,

$$\text{Consistency} + \text{stability} \Rightarrow \text{Convergence.} \qquad (9.1)$$

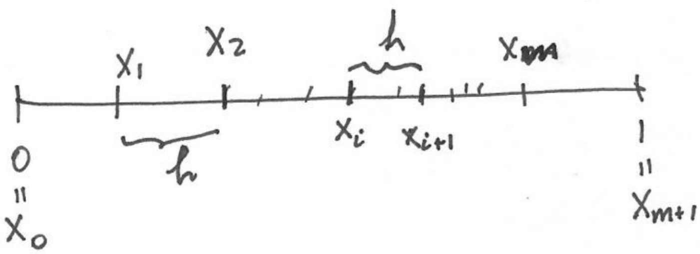
Notice that the size of A^h grows when $h \to 0$
Therefore, proving stability may be hard in general.

The concept of stability depends on the BVP or IBVP analyzed. The goal is that given a finite difference method for a particular problem that the definition of stability allows one to prove convergence using (9.1) and that possible to verify it for the FDM studied.

Summary

Consistency - Convergence - stability (2nd class)

Global error $= \; E = \underset{\text{Approx - Exact}}{\vec{U} - \hat{U}} = \begin{bmatrix} v_1 - \overset{=u_1}{u(x_1)} \\ \vdots \\ v_m - \overset{=u_m}{u(x_m)} \end{bmatrix}$



uniform grid.

Two norms to measure global error:

$$\|E\|_\infty = \max_{1\le j\le m} |v_j - u_j|$$

and

$$\|E\|_2 = \left[h \sum |v_j - u_j|^2 \right]^{1/2}.$$

Convergence:

$$\|E^h\| \xrightarrow[h\to 0]{} 0$$

Consistency and Local Truncation error:

Cont. Equ. : $(Pu)(x) \equiv u''(x) - f(x) = 0$ + BC's.

Discrete Equ: $P_\Delta v_j \equiv \dfrac{v_{j+1} - 2v_j + v_{j-1}}{h^2} - f(x_j) = 0$

$j = 1, 2, \ldots, m$

Local Truncation Error: $P_\Delta u_j - (Pu)(x_j) = \tau_j$, $j = 1, 2, \ldots, m$

$$\vec{r}^h = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix}$$

Two matrix equations:

$$A\vec{U}^h = \vec{F}^h$$

$$A\hat{U}^h = \vec{F}^h + \vec{r}^h$$

If $\|\vec{r}^h\| \xrightarrow[h\to 0]{} 0$ the numerical method is consistent.

<u>Stability</u>:

Numerical Method : $A^h \vec{U}^h = \vec{F}^h$

Stable if

i) $(A^h)^{-1}$ exists for all $h < h_0$

It means the discrete equ. of the num. method has a unique solution for all $h < h_0$.

ii) $\|(A^h)^{-1}\| \leq C$, for $h < h_0$.

<u>Matrix Norms:</u>

If $A_{n\times n}$ matrix is such that

$$\|A\vec{x}\| \leq C\|\vec{x}\|, \text{ for all } x \in \mathbb{R}^n \quad (*)$$

$\|A\| =$ Smallest C Satisfying $(*)$

Equivalent to say

$$\|A\| = \max_{\substack{\vec{x}\in\mathbb{R}^n \\ \vec{x}\neq 0}} \frac{\|A\vec{x}\|}{\|\vec{x}\|} = \max_{\substack{\vec{x}\in\mathbb{R}^n \\ \|\vec{x}\|=1}} \|A\vec{x}\|.$$

Particular cases:

$$\|A\|_\infty = \max_{1 \le j \le m} \left(\sum_{j=1}^{n} |a_{ij}| \right) \quad (\text{max row sum})$$

or

$$\|A\|_\infty = \max_{1 \le i \le n} \left(\sum_{j=1}^{n} |a_{ij}| \right) \quad (\text{max Column Sum})$$

and

$$\|A\|_2 = \sqrt{\rho(A^T A)}.$$

$\rho(B)$: spectral radius of B
or max modulus of an eigenvalue.

Fundamental Thm.:

$$A^h \vec{U}^h = \vec{F}^h \quad \text{and} \quad A^h \hat{U}^h = \vec{F}^h + \vec{\tau}^h$$

then

$$A^h \vec{E}^h = A(\vec{U}^h - \hat{U}^h) = -\vec{\tau}^h$$

If $(A^h)^{-1}$ exists and $\|(A^h)^{-1}\| \le C$

Then

$$\vec{E}^h = -(A^h)^{-1} \vec{\tau}^h$$

and

$$\|\vec{E}^h\| \le \|(A^h)^{-1}\| \, \|\vec{\tau}^h\| \overset{\text{Stability}}{\le} C \, \|\vec{\tau}^h\|$$

Now, the numerical method is consistent

then $\|\vec{\tau}^h\| \xrightarrow[h \to 0]{} 0 \implies \|\vec{E}^h\| \xrightarrow[h \to 0]{} 0$ (convergence)

Furthermore, $\|\vec{E}^h\| = O(h^p)$ if $\|\vec{\tau}^h\| = O(h^p)$.