Numerical Solution of Partial Differential Equations

Chapter 1 Piecewise Polynomials and Approximation Theory

§1.1 Polynomials

§1.2 Piecewise Polynomials on Fixed Partition

§1.3 ReLU Neural Network

Chapter 2 Numerical Methods for PDEs

§2.1 Elliptic PDEs

§2.1.1 Equivalent Formulations (Primal, Dual, LS)

§2.1.2 Numerical Methods

§2.2 Hyperbolic Conservation Laws (HCLs)

§2.2.1 Linear Advection–Reaction Problem

§2.2.2 Nonlinear scalar HCLs

§2.3 Elliptic System

§2.3.1 Linear elasticity

§2.3.2 Stoke's equation

§2.3.3 Maxwell's equation

Chapter 3 Linear and Nonlinear Solvers

§3.1 Linear solvers

§3.2 Nonlinear solvers

Chapter 1 Piecewise Polynomials and Approximation Theory

§1.1 Polynomials

- polynomials of degree $\le n$

$$\mathcal{P}_n = \left\{ c_0 + c_1 x + \cdots + c_n x^n \;\middle|\; c_i \in R \right\}$$

$$= \text{span}\left\{1, x, \cdots, x^n\right\} = \text{span}\left\{L_0(x), L_1(x), \cdots, L_n(x)\right\}$$

where $L_i(x) \in \mathcal{P}_i$ and $\{L_i(x)\}_{i=0}^n$ is an orthonormal basis with respect to an inner product.

- approximation problems

(1) Taylor's approximation at x_0

Problem Find $T_n(x) \in \mathcal{P}_n$ such that $T_n^{(i)}(x_0) = f^{(i)}(x_0)$ for $i = 0, 1, \cdots, n$

solution $T_n(x) = f(x_0) + f'(x_0)(x - x_0) + \cdots + \dfrac{f^{(n)}(x_0)}{n!}(x - x_0)^n$

the error estimate $e_T(x) = f(x) - T_n(x) = \dfrac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^n$, $\exists\, \xi$ between x and x_0

Proof $e_T^{(i)}(x_0) = 0$ for $i = 0, 1, \cdots, n$

$\Longrightarrow e_T(x) = R(x)(x - x_0)^{n+1}$

Set $g(t) = e_T(t) - R(x)(t - x_0)^{n+1} \Rightarrow g^{(i)}(x_0) = 0$ for $i = 0, 1, \cdots, n$ and $g(x) = 0$

$\xRightarrow{\text{Rolle's Theorem}} \exists\, \xi$ between x and x_0 s.t. $0 = g^{(n+1)}(\xi) = f^{(n+1)}(\xi) - R(x)(n+1)!$

$\Longrightarrow R(x) = f^{(n+1)}(\xi) / (n+1)!$ #

(2) Interpolation at $x_0 < x_1 < \cdots < x_n$

Problem Find $f_I(x) \in \mathcal{P}_n$ s.t. $f_I(x_i) = f(x_i)$ for $i = 0, 1, \cdots, n$.

solution I Let $f_I(x) = \sum_{j=0}^{n} c_j x^j$, then

$$\underbrace{\begin{pmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^n \end{pmatrix}}_{V} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix} \Longrightarrow \underset{\text{Vandermonde matrix}}{V} \vec{c} = \vec{F}$$

Lemma $\det V = \prod_{0 \le i < j \le n} (x_j - x_i)$.

Solution II

$$f_I(x) = \sum_{i=0}^{n} f(x_i)\, l_i(x)$$

where $l_i(x) = \prod_{\substack{j=0 \\ j \ne i}}^{n} \frac{x - x_j}{x_i - x_j}$ is the Lagrange nodal basis function,

i.e., $l_i(x_j) = \delta_{ij} = \begin{cases} 1, & j = i \\ 0, & j \ne i \end{cases}$.

the error estimate

$$e_I(x) = f(x) - f_I(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{j=0}^{n} (x - x_j)$$

Proof $e_I(x_i) = 0$ for $i = 0, 1, \cdots, n \Longrightarrow e_I(x) = R(x) \prod_{j=0}^{n} (x - x_j)$

Set $g(t) = e_I(t) - R(x) \prod_{j=0}^{n} (t - x_j) \Longrightarrow g(x) = 0$ and $g(x_i) = 0$ for $i = 0, 1, \cdots, n$

$\overset{\text{Rolle's Thrm}}{\Longrightarrow} \exists\, \xi$ s.t. $0 = g^{(n+1)}(\xi) = f^{(n+1)}(\xi) - R(x)(n+1)!$

#

(3) <u>Least-squares Approximation</u>

Let $f(x)$ be a given function defined on interval $I=[a,b]$ and $w(x)\geq 0$ be a weight function. Denote the weighted $L^2(I)$ inner product and its induced norm, respectively, by

$$\langle f, g\rangle = \int_a^b w(x)\, f(x) g(x)\,dx \quad \text{and} \quad \|f\|^2 = \langle f, f\rangle.$$

<u>Problem</u> Find $p_n(x)\in\mathcal{P}_n$ s.t.

$$(LS) \qquad \|f-p_n\| = \min_{q\in\mathcal{P}_n} \|f-q\| \iff \|f-p_n\|\leq\|f-q\| \ \ \forall q\in\mathcal{P}_n.$$

<u>Lemma</u> $p_n=\arg\min_{q\in\mathcal{P}_n}\|f-q\| \iff \left[\begin{array}{l}\text{Find } p_n\in\mathcal{P}_n \text{ s.t.}\\ (p_n, q) = (f,q) \ \ \forall q\in\mathcal{P}_n. \quad (VP)\end{array}\right.$

<u>Proof</u> "$\Rightarrow$" Let $p_n\in\mathcal{P}_n$ be the solution of (LS). For any $q\in\mathcal{P}_n$ and $t\in\mathbb{R}$, set $g(t)=\|f-(p_n+tq)\|^2 = \|f-p_n\|^2 - 2t(f-p_n, q)+t^2\|q\|^2$

$p_n+tq\in\mathcal{P}_n \implies g(t)$ has a minimum at $t=0$

$\implies 0=g'(0)=-(f-p_n,q) \implies (p_n,q)=(f,q).$

"$\Leftarrow$" Let $p_n\in\mathcal{P}_n$ be the solution of (VP)

$\implies (f-p_n, q)=0 \ \ \forall q\in\mathcal{P}_n$

$\implies g(t)=\|f-p_n\|^2+t^2\|q\|^2 \geq g(0)=\|f-p_n\|^2$

or $\forall q\in\mathcal{P}_n$

$$\|f-q\|^2 = \|(f-p_n)+(p_n-q)\|^2 = \|f-p_n\|^2+\|p_n-q\|^2+2\overbrace{(f-p_n, \underbrace{p_n-q}_{\in\mathcal{P}_n})}^{0}$$
$$\geq \|f-p_n\|^2. \qquad \#$$

Solution I based on (VP) Let $p_n(x) = \sum_{i=0}^{n} c_i \varphi_i(x)$

where $\{\varphi_i(x)\}_{i=0}^{n}$ is any basis functions of $\mathcal{P}_n$.

$$\forall\, j = 0, 1, \cdots, n, \quad (VP) \Longrightarrow (f, \varphi_j) = (p_n, \varphi_j) = \left(\sum_{i=0}^{n} c_i \varphi_i, \varphi_j\right)$$

$$\left((\varphi_0, \varphi_j), (\varphi_1, \varphi_j), \cdots, (\varphi_n, \varphi_j)\right) \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix} = \sum_{i=0}^{n} c_i (\varphi_i, \varphi_j)$$

$$\Longrightarrow \underbrace{\begin{pmatrix} (\varphi_0, \varphi_0) & (\varphi_1, \varphi_0) & \cdots & (\varphi_n, \varphi_0) \\ (\varphi_0, \varphi_1) & (\varphi_1, \varphi_1) & \cdots & (\varphi_n, \varphi_0) \\ \vdots & \vdots & & \vdots \\ (\varphi_0, \varphi_n) & (\varphi_1, \varphi_n) & \cdots & (\varphi_n, \varphi_n) \end{pmatrix}}_{M} \vec{c} = \underbrace{\begin{pmatrix} (f, \varphi_0) \\ (f, \varphi_1) \\ \vdots \\ (f, \varphi_n) \end{pmatrix}}_{\vec{f}} \Longrightarrow M\vec{c} = \vec{f}$$

$$\Longrightarrow \vec{c} = M^{-1} \vec{f}.$$

Lemma M is symmetric, positive definite.

Proof $m_{ij} = (\varphi_j, \varphi_i) = (\varphi_i, \varphi_j) = m_{ji} \Longrightarrow$ M is symmetric.

$\forall\, \vec{\xi} = (\xi_0, \xi_1, \cdots, \xi_n)^T \in R^{n+1}$, let $g(x) = \sum_{i=0}^{n} \xi_i \varphi_i(x)$

$$\vec{\xi}^T M \vec{\xi} = \vec{\xi}^T \left(\sum_{i=0}^{n} \xi_i (\varphi_i, \varphi_0), \sum_i \xi_i (\varphi_i, \varphi_1), \cdots, \sum_i \xi_i (\varphi_i, \varphi_n)\right)^T$$

$$= \vec{\xi}^T \left((g, \varphi_0), (g, \varphi_1), \cdots, (g, \varphi_n)\right)^T$$

$$= \sum_{i=0}^{n} \xi_i (g, \varphi_i) = (g, g) = \|g\|^2 > 0$$

$\Longleftrightarrow$ g is not identically zero $\Longleftrightarrow$ $\vec{\xi} \neq \vec{0}$. #

<u>Solution II</u> Let $\{L_i(x)\}_{i=0}^{n}$ be ~~a~~ basis functions of $\mathcal{P}_n$ and they are orthogonal, i.e., $(L_i, L_j) = \delta_{ij} = \begin{cases} 1, & i=j \\ 0, & i\neq j \end{cases}$.

$$\Longrightarrow \quad P_n(x) = \sum_{i=0}^{n} (f, L_i)\, L_i(x).$$

<u>Proof</u> $\{L_i(x)\}_{i=0}^{n}$ is orthonormal $\Longrightarrow M = I$

$$\Rightarrow \quad c_i = (f, L_i) \quad \text{for } i=0,1,\cdots,n. \qquad \#$$

<u>the error estimate</u>

(1) Assume that $\{L_i(x)\}_{i=0}^{\infty}$ is a orthonormal basis of $L^2(I)$ w.r.t. the weight inner product and that $L_i(x) \in \mathcal{P}_i$.

$$\Longrightarrow \quad f(x) = \sum_{i=0}^{\infty} (f, L_i)\, L_i(x)$$

$$\Longrightarrow \quad f(x) - P_n(x) = \sum_{i=n+1}^{\infty} (f, L_i)\, L_i(x)$$

(2) $\|f-P_n\|^2 = (f-P_n, f-P_n) = (f-P_n, f-v) \quad \forall\, v \in \mathcal{P}_n$

$$\leq \|f-P_n\|\,\|f-v\|$$

$$\Longrightarrow \quad \|f-P_n\| \leq \|f-v\| \quad \forall\, v\in \mathcal{P}_n$$

$$\Longrightarrow \quad \|f-P_n\| \leq \|f-f_I\|$$ where $f_I(x)$ ~~is~~ an interpolant of f w.r.t. any partition $a \leq x_0 < x_1 < \cdots < x_n \leq b$.

§1.2 Piecewise Polynomials on Fixed Partition

§1.2.1 One dimension

Let $\Delta: a = x_0 < x_1 < \cdots < x_n = b$ be a partition of the interval $I = [a, b]$.

Set $I_i = [x_{i-1}, x_i]$, $h_i = x_i - x_{i-1}$, for $i = 1, \cdots, n$.

<u>Spline</u> $S_m^k(\Delta) = \left\{ v(x) \in C^k(I) \;\middle|\; v\big|_{I_i} \in \mathcal{P}_m \text{ for } i = 1, \cdots, n \right\}$.

<u>C^0-finite element</u> $(k = 0)$

$$S_m^0(\Delta) = \left\{ v(x) \in C^0(I) \;\middle|\; v\big|_{I_i} \in \mathcal{P}_m \text{ for } i = 1, \cdots, n \right\}.$$

<u>dimension</u> $\dim S_m^0(\Delta) = \underbrace{(m+1)}_{\dim \mathcal{P}_m} \underbrace{n}_{\#\text{ of intervals}} - \underbrace{(n-1)}_{\#\text{ of continuity constraints}} = mn + 1$

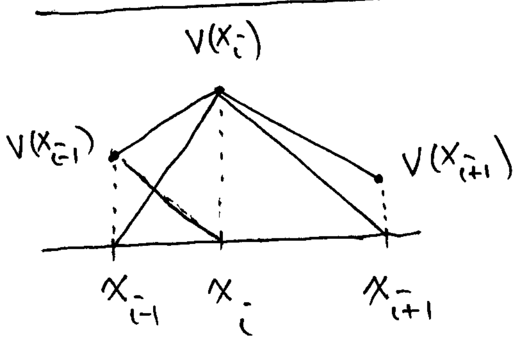
degrees of freedom ? basis functions ?

<u>linear element</u> $(m = 1)$ $v(x) \in S_1^0(\Delta)$, $v(x) = ?$

$? = v(x) \in C^0(I) \Longrightarrow v(x_i^-) = v(x_i^+)$ for $i = 1, \cdots, n-1$

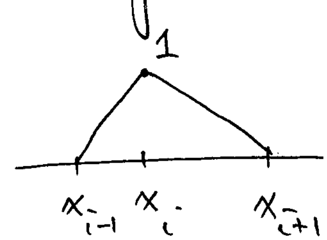
$\Longrightarrow \{ v(x_j) \}_{j=0}^{n}$ — nodal ~~variables~~ values (variables)

<u>On $I_i = (x_{i-1}, x_i)$</u> $v(x) = v(x_{i-1}) \varphi_{i-1}(x) + v(x_i) \varphi_i(x)$

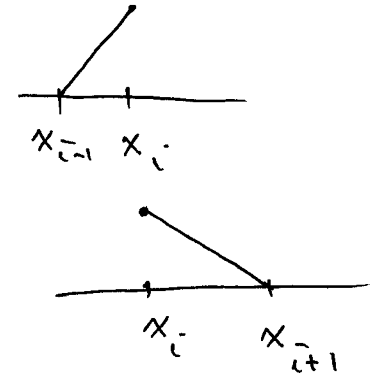


$\{ \varphi_i(x) \}_{i=0}^{n}$ — nodal basis function

$\varphi_i(x_j) = \delta_{ij}$



$$\varphi_i(x) = \begin{cases} (x - x_{i-1})/(x_i - x_{i-1}), & x \in I_i \\ (x_{i+1} - x)/(x_{i+1} - x_i), & x \in I_{i+1} \\ 0, & x \notin I_i \cup I_{i+1} \end{cases}$$



$$\Longrightarrow \quad v(x) = \sum_{i=0}^{n} v(x_i)\, \varphi_i(x)$$

The triple $(I, \mathcal{P}, \mathcal{N})$ is called a finite element

where, e.g., $I = I_i = [x_{i-1}, x_i]$

$\mathcal{P} = \mathcal{P}_1 = \text{span}\{\varphi_{i-1}(x), \varphi_i(x)\}$, and $\mathcal{N} = \{N_{i-1}, N_i\}$ with $N_{i-1}(v) = v(x_{i-1})$, $N_i(v) = v(x_i)$

<u>Definition</u> $\mathcal{N}$ determines $\mathcal{P}$ $\Longleftrightarrow$ "$\psi \in \mathcal{P}$ with $N(\psi) = 0 \ \forall N \in \mathcal{N}$ $\Longrightarrow \psi \equiv 0$"

<u>Lemma</u> $\{N_{i-1}, N_i\}$ determines $\mathcal{P}_1$.

<u>Proof</u> $\forall \psi \in \mathcal{P}_1$, assume that $\psi(x_{i-1}) = 0$ and $\psi(x_i) = 0$

$$\Longrightarrow \psi(x) = a + bx \quad \text{and} \quad \begin{cases} a + b x_{i-1} = 0 \\ a + b x_i = 0 \end{cases} \Longrightarrow \begin{pmatrix} 1 & x_{i-1} \\ 1 & x_i \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Longrightarrow a = b = 0 \Longrightarrow \psi(x) \equiv 0 \ \text{in } I_i.$$

<u>the error estimate</u> let $f_I(x) = \sum_{i=0}^{n} f(x_i)\, \varphi_i(x)$

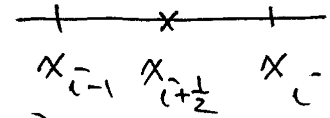
On $I_i = [x_{i-1}, x_i]$, $f_I(x) = f(x_{i-1})\, \varphi_{i-1}(x) + f(x_i)\, \varphi_i(x)$

$$f(x) - f_I(x) = \frac{f''(\xi_i)}{2} (x - x_{i-1})(x - x_i)$$

$$\max_{x \in I_i} |f(x) - f_I(x)| \le \frac{1}{2} \max_{x \in I_i} |f''(\xi_i)| \max_{x \in I_i} |(x - x_{i-1})(x - x_i)| = \frac{1}{8} h_i^2 \|f''\|_{\infty, I_i}$$

quadratic element $(m=2)$ $\quad (I_i, \mathcal{P}_2, \mathcal{N})$

$I_i = [x_{i-1}, x_i]$, $\mathcal{N} = \{N_{i-1}, N_i, \hat{N}_i\}$

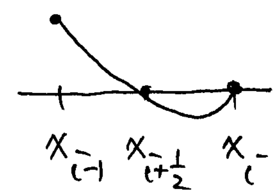


$C^0(I) \Longrightarrow N_{i-1}(v) = v(x_{i-1})$ and $N_i(v) = v(x_i)$

choose $\hat{N}_i(v) = v(x_{i+\frac{1}{2}})$ where $x_{i+\frac{1}{2}} = \frac{1}{2}(x_{i-1} + x_i)$.

nodal basis functions

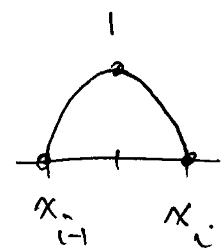
$\underline{\varphi_{i-1}(x)} \in \mathcal{P}_2$: $\varphi_{i-1}(x_{i-1}) = 1$, $\varphi_{i-1}(x_{i+\frac{1}{2}}) = 0$, $\varphi_{i-1}(x_i) = 0$



$$\Longrightarrow \varphi_{i-1}(x) = a(x - x_{i+\frac{1}{2}})(x - x_i)$$

$$1 = \varphi_{i-1}(x_{i-1}) = a\,\frac{h_i}{2}\,h_i \Longrightarrow \varphi_{i-1}(x) = \frac{2}{h_i^2}(x - x_{i+\frac{1}{2}})(x - x_i)$$

similarly, $\varphi_i(x) = \frac{2}{h_i^2}(x - x_{i-1})(x - x_{i+\frac{1}{2}})$



$\underline{\varphi_{i+\frac{1}{2}}(x)} \in \mathcal{P}_2$: $\varphi_{i+\frac{1}{2}}(x_{i-1}) = \varphi_{i+\frac{1}{2}}(x_i) = 0$, $\varphi_{i+\frac{1}{2}}(x_{i+\frac{1}{2}}) = 1$

$$\varphi_{i+\frac{1}{2}}(x) = a(x - x_{i-1})(x - x_i)$$

$$\Longrightarrow \varphi_{i+\frac{1}{2}}(x) = -\frac{4}{h_i^2}(x - x_{i-1})(x - x_i)$$

$\boxed{\cancel{v \in S_2^0(\Delta) \Longleftrightarrow v(x)}}$

$$v \in S_2^0(\Delta) \Longrightarrow v(x) = \sum_{i=0}^{n} v(x_i)\varphi_i(x) + \sum_{i=1}^{n-1} v(x_{i+\frac{1}{2}})\varphi_{i+\frac{1}{2}}(x)$$