§1.2 One dimensional piecewise polynomials

§1.2.1 Fixed Mesh $\Delta:\ a=x_0<x_1<\cdots<x_n=b$

$$S_m^k(\Delta)=\left\{v\in C^k(a,b)\ \middle|\ v|_{[x_{i-1},x_i]}\in \mathcal{P}_m \text{ for } i=1,\cdots,n\right\}$$

§1.2.2 Moving Mesh (Free-knot spline and Neural Network

Example (non-smooth function)

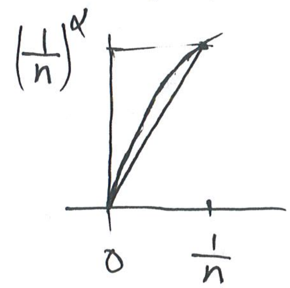
$f(x)=x^\alpha$ with $0<\alpha<1$ defined on $[0,1]$

(1) C^0-piecewise linear approximation on a uniform mesh $\left\{x_i=\frac{i}{n}\right\}_{i=0}^n$.

$$f_I(x)=\sum_{i=0}^n f(x_i)\,\varphi_i(x)=\sum_{i=1}^n x_i^\alpha\,\varphi_i(x)$$

On $I_1=[x_0,x_1]=[0,\frac{1}{n}]$: $\quad f_I(x)=x_1^\alpha\varphi_1(x)=\left(\frac{1}{n}\right)^\alpha\cdot\frac{x}{\frac{1}{n}}$

$$\Rightarrow e_I(x)=f(x)-f_I(x)=x^\alpha-n^{1-\alpha}x\ \cancel{=x^\alpha\left[1-(nx)^{1-\alpha}\right]}$$



On $[\varepsilon,\frac{1}{n}]$ $\quad \max\limits_{x\in[\varepsilon,\frac{1}{n}]}|e_I(x)|=e_I\left(\frac{1}{n}\cdot\alpha^{\frac{1}{1-\alpha}}\right)$

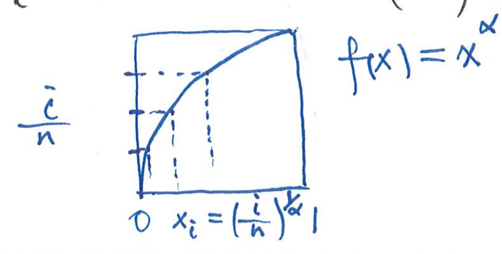
$$\max_{x\in[0,\frac{1}{n}]}|e_I(x)|=\left(\frac{1}{n}\right)^\alpha\left[\alpha^{\frac{\alpha}{1-\alpha}}-\alpha^{\frac{1}{1-\alpha}}\right]=\left(\frac{1}{n}\right)^\alpha\frac{1-\alpha}{\alpha}\cdot\alpha^{\frac{1}{1-\alpha}}$$

$$=\left(\frac{1}{n}\right)^\alpha(1-\alpha)\,\alpha^{\frac{\alpha}{1-\alpha}}$$

(2) C^0-piecewise linear approximation on a specially choosen mesh $\{x_i\}_{i=0}^n$

where $f(x_i)=\frac{i}{n}\ \Rightarrow\ x_i^\alpha=\frac{i}{n}\ \Rightarrow\ x_i=\left(\frac{i}{n}\right)^{\frac{1}{\alpha}}$.

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



On $[x_0, x_1] = [0, (\frac{1}{n})^{1/\alpha}] = (\frac{1}{n})^{1/\alpha}[0,1]$

$$e_I(x) = f(x) - f_I(x) = x^\alpha - x_1^\alpha \cdot \frac{x}{x_1} = x^\alpha - \left(\frac{1}{n}\right)^{\frac{\alpha-1}{\alpha}} x$$

let $\hat{x} \neq 0$ and $e_I'(\hat{x}) = 0 \Rightarrow \hat{x} = \alpha^{\frac{1}{1-\alpha}} \left(\frac{1}{n}\right)^{\frac{1}{\alpha}}$

$$\Rightarrow \max_{x\in[x_0,x_1]} e_I(x) = e_I(\hat{x}) = \alpha^{\frac{1}{1-\alpha}}\left(\frac{1}{n}\right) - \left(\frac{1}{n}\right)^{\frac{\alpha-1}{\alpha}} \cdot \alpha^{\frac{1}{1-\alpha}}\left(\frac{1}{n}\right)^{\frac{1}{\alpha}}$$

$$= \frac{1}{n}(1-\alpha)\,\alpha^{\frac{\alpha}{1-\alpha}}$$

On $[x_{n-1}, x_n] = \left[\left(\frac{n-1}{n}\right)^{\frac{1}{\alpha}}, 1\right]$

$$? \geq \max_{x\in[x_{n-1},x_n]} |e_I(x)|$$

<u>Free-knot spline</u> (1981 Schumaker "Spline Function: Basic Theory")

$$S_{1,n} = \left\{ s(x) \in C^0[a,b] \;\middle|\; s(x) \text{ is piecewise linear with } n \text{ knots in } [a,b] \right\}$$

Let $a \leq x_1 < x_2 < \cdots < x_n \leq b$ and

$$\varphi_i(x; x_{i-1}, x_i, x_{i+1}) = \begin{cases} (x - x_i)/(x_{i-1} - x_i), & x \in I_i = [x_{i-1}, x_i] \\ (x - x_{i-1})/(x_i - x_{i-1}), & x \in I_{i+1} \\ 0, & \text{otherwise} \end{cases}$$

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

$$\varphi_1(x) = \begin{cases} \frac{x - x_2}{x_1 - x_2}, & x \in [a, x_2] \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad \varphi_n(x) = \begin{cases} 0, & x \in [a, x_{n-1}] \\ (x - x_{n-1})/(x_n - x_{n-1}), & x \in I_n \end{cases}$$

$$s(x) \in S_{1,n} \Longrightarrow s(x) = \sum_{i=1}^{n} c_i \varphi_i (x; x_{i-1}, x_i, x_{i+1})$$

<u>parameters</u>
- linear $\{c_i\}_{i=1}^{n}$
- nonlinear $\{x_i\}_{i=1}^{n}$

<u>Remark</u> $S_{1,n}$ is not a linear space

<u>Proof</u> Let $s_1(x),\ s_2(x) \in S_{1,n}$ with knots $\{x_i^1\}_{i=1}^{n}$ and $\{x_i^2\}_{i=1}^{n}$

If $\{x_i^1\}$ and $\{x_i^2\}$ have different knot

$$\Longrightarrow s_1(x) + s_2(x) \notin S_{1,n}, \text{ but } \boxed{s_1(x) + s_2(x) \in S_{1,2n}}$$

Let $f_I(x) \in S_{n,1}$ be the interpolant of $f(x)$

$$\vec{x} = (x_1, \cdots, x_n)$$

$$\Longrightarrow f_I(x; \vec{x}) = \sum_{i=1}^{n} f(x_i)\, \varphi(x; x_{i-1}, x_i, x_{i+1})$$

Find $\vec{x}$ such that $a \le x_1 < \cdots < x_n \le b$ which minimizes

$$\left\| f - f_I(x; \vec{x}) \right\| = \min_{\vec{y} \in \mathcal{X}_n = \{\vec{x} \mid a \le x_1 < \cdots < x_n \le b\}} \| f(x) - f_I(x; \vec{y}) \|$$

- Free-knot spline

$$S^0_{1,n} = \left\{ s(x) \in C^0(a,b) \;\middle|\; \begin{array}{l} s(x) \text{ is piecewise linear with} \\ n \text{ knots } a \le \alpha_1 < \alpha_2 < \cdots < \alpha_n \le b \end{array} \right\}$$

$$\forall\, s(x) \in S^0_{1,n}, \quad \exists\, \{\alpha_i\}_{i=1}^n \in X_n = \left\{ \{\alpha_i\}_{i=1}^n \;\middle|\; a \le \alpha_1 < \cdots < \alpha_n \le b \right\}$$

s.t. $\displaystyle s(x) = \sum_{i=1}^n c_i \varphi_i(x; \vec{\alpha})$

where $\varphi_i(x;\vec{\alpha})$ is the nodal basis function based on $\vec{\alpha} = \{\alpha_i\}_{i=1}^n$.

$$\Rightarrow S^0_{1,n} = \left\{ \sum_{i=1}^n c_i \varphi_i(x;\vec{\alpha}) \;\middle|\; c_i \in R \text{ and } \vec{\alpha} \in X_n \right\}.$$

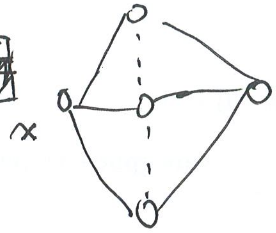
Least-squares Approximations Given $f(x)$ defined on $[a,b]$,

(1) find $f_n(x) \in S^0_{1,n}$ s.t. $\displaystyle \|f - f_n\| = \min_{s \in S^0_{1,n}} \|f - s\|$;

(2) let $\displaystyle f_I(x;\vec{\alpha}) = \sum_{i=1}^n f(\alpha_i)\varphi_i(x;\vec{\alpha})$,

find $\vec{\alpha}^* \in X_n$ s.t. $\displaystyle \vec{\alpha}^* = \operatorname*{argmin}_{\vec{\alpha} \in X_n} \|f - f_I(\cdot;\vec{\alpha})\|$

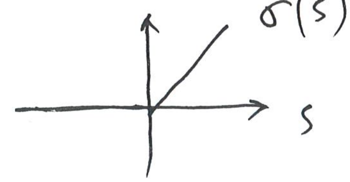
§1.2.3 ReLU neural network



$$N(x) = \sum_{i=1}^{n} c_i\,\sigma(w_i x - b_i) + c_0$$

w_i — weight, b_i — bias

$\sigma(s) = \max\{s, 0\}$

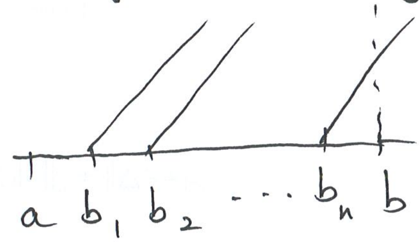


normalization

$$|w_i| = 1 \implies w_i = \pm 1$$

$$\mathcal{N}_n = \left\{ \sum_{i=1}^{n} c_i\,\sigma(x - b_i) + c_0 \;\middle|\; c_i \in \mathbb{R},\ a \le b_1 < \cdots < b_n \le b \right\}$$

Lemma $\{\sigma(x-b_i)\}_{i=1}^{n}$ is linearly independent if $b_i \ne b_j$ for all $i \ne j$.



Proof $\displaystyle\sum_{i=1}^{n} c_i\,\sigma(x-b_i) = 0 \quad \forall\, x \in [a,b]$

$$\implies 0 = c_1\,\sigma(x-b_1) \quad \forall\, x \in [b_1, b_2] \implies c_1 = 0$$

$$\implies 0 = \sum_{i=2}^{n} c_i\,\sigma(x-b_i) \quad \forall\, x \in [b_2, b] \implies \cdots \implies c_i = 0 \quad \forall i. \qquad \#$$

~~Least-squares approximation~~

~~find $f_n(x)$~~

Relation between $S_{1,n}^{0}$ and $\mathcal{N}_n$.

- produce the same class of continuous piecewise linear functions with moving mesh
- local and global basis functions.

Least-squares approximation

(C.-Dokotorova-Falgout-Herrera (1-D)
C.-Ding-Liu-Liu-Xia (d-D))

find $f_n(x) \in \mathcal{N}_n$ s.t. $\|f-f_n\| = \min_{s\in\mathcal{N}_n} \|f-s\|$

Let $f_n(x) = f_n(x;\vec{c},\vec{b}) = c_0 + \sum_{i=1}^{n} c_i\,\sigma(x-b_i)$, $\vec{c}=(c_0,c_1,\cdots,c_n)$, $\vec{b}=(b_1,\cdots,b_n)$

$$J(f_n) = \frac{1}{2}\|f-f_n\|^2 = \frac{1}{2}\int_a^b (f-f_n)^2(x)\,dx.$$

Optimality conditions

$$\nabla_{\vec{c}} J(f_n) = \nabla_{\vec{c}}\,\frac{1}{2}\int_a^b (f-f_n)^2 dx = \frac{1}{2}\int_a^b \nabla_{\vec{c}}(f-f_n)^2 = -\int_a^b (f-f_n)\,\nabla_{\vec{c}} f_n$$

$$\nabla_{\vec{c}} f_n = (1, \sigma(x-b_1), \cdots, \sigma(x-b_n))^T \equiv \vec{\Sigma}(x;\vec{b}) \quad \text{and } f_n = \vec{c}^{\,T}\vec{\Sigma}$$

$$\Longrightarrow \nabla_{\vec{c}} J(f_n) = -\int_a^b f\,\vec{\Sigma}\,dx + \int_a^b \left(\vec{c}^{\,T}\vec{\Sigma}\right)\vec{\Sigma}\,dx = \left(\int_a^b \vec{\Sigma}\,\vec{\Sigma}^T dx\right)\vec{c} - \int_a^b f\,\vec{\Sigma}\,dx$$

$$(m_{ij})_{(n+1)\times(n+1)} = M(\vec{b}) = \int_a^b \vec{\Sigma}\,\vec{\Sigma}^T dx, \quad \vec{F}(\vec{b}) = \int_a^b f\,\vec{\Sigma} = (f_i)_{(n+1)\times 1}$$

$$m_{ij}(\vec{b}) = \int_a^b \varphi_i(x)\,\varphi_j(x)\,dx \quad \text{where } \varphi_0(x)=1 \text{ and } \varphi_i(x)=\sigma(x-b_i)$$

$$f_i(\vec{b}) = \int_a^b f(x)\,\varphi_i(x)\,dx.$$

$$(1) \qquad \boxed{0 = \nabla_{\vec{c}} J(f_n) = M(\vec{b})\,\vec{c} - \vec{F}(\vec{b})}$$

Lemma $M(\vec{b})$ is symmetric, positive definite.

$$\nabla_{\vec b} J(u_n) = \frac{1}{2}\nabla_{\vec b}\int_a^b (f-f_n)^2\,dx \overset{?}{=} \int_a^b (f-f_n)\,\nabla_{\vec b} f_n\,dx$$

$$\nabla_{\vec b} f_n = \left(\frac{\partial f_n}{\partial b_1},\dots,\frac{\partial f_n}{\partial b_n}\right)^T = \left(c_1\frac{\partial \sigma(x-b_1)}{\partial b_1},\dots,c_n\frac{\partial \sigma(x-b_n)}{\partial b_n}\right)^T$$

$$= -\big(c_1 H(x-b_1),\dots,c_n H(x-b_n)\big)^T = D(\vec c)\,\vec H(x;\vec b)$$

$$\boxed{\begin{array}{l} H(x)=\sigma'(x)\\ \text{weak derivative}\end{array}}$$

where $D(\vec c)=\mathrm{diag}(c_1,\dots,c_n)$ and $\vec H(x;\vec b) = \big(H(x-\sigma_1),\dots,H(x-b_n)\big)^T$.

$$(2)\qquad \boxed{0=\nabla_{\vec b} J(u_n) = D(\vec c)\int_a^b (f-f_n)\,\vec H\,dx} \xRightarrow[c_i\neq 0]{} \int_a^b (f-f_n)\,\vec H = 0$$

Hessian Matrix

$$H_i(x)\equiv H(x-b_i)$$

$$\frac{\partial}{\partial b_i}\int_a^b (f-f_n)\,H(x-b_j)\,dx \overset{?}{=} \int_a^b \frac{\partial}{\partial b_i}\Big[(f-f_n)\,H(x-b_j)\Big]dx$$

$$= -\int_a^b \frac{\partial f_n}{\partial b_i}H_j + \int_a^b (f-f_n)\frac{\partial}{\partial b_i}H_j$$

$$= c_i\int_a^b H_i H_j - \delta_{ij}\int_a^b (f-f_n)\,\delta(x-b_j)$$

$$= c_i\int_a^b H_i H_j + \delta_{ij}\big(f_n(b_j)-f(b_j)\big)$$

$$\boxed{\begin{array}{l}\dfrac{\partial H_j}{\partial b_i} = \begin{cases}0, & i\neq j\\ -\delta(x-b_j), & i=j\end{cases}\\[2ex] \delta_{ij} = \begin{cases}0, & i\neq j\\ 1, & i=j\end{cases}\end{array}}$$

$$\nabla^2_{\vec b} J(u_n) = \nabla_{\vec b}\big(\nabla_{\vec b} J(u_n)\big)^T = \nabla_{\vec b}\int_a^b (f-f_n)(H_1,\dots,H_n)\,dx\; D(\vec c)$$

$$= \left(\frac{\partial}{\partial b_i}\int_a^b (f-f_n)H_j\right)_{n\times n} = \left[D(\vec c)\left(\int_a^b H_i H_j\right)_{n\times n} + \mathrm{diag}\big((f_n-f)(b_1),\dots,(f_n-f)(b_n)\big)\right]D(\vec c)$$

$$= D(\vec c)\underbrace{\int_a^b \vec H\vec H^T dx}_{A(\vec b)}\,D(\vec c) + D(\vec w)\,D(\vec c), \qquad D(\vec w) = \mathrm{diag}(w_1,\dots,w_n)$$

$$w_i = (f_n-f)(b_i)$$

Leibniz Integral Rule

$$\frac{d}{dx}\int_{a(x)}^{b(x)} f(x,t)\,dt = f(x,b(x))\,b'(x) - f(x,a(x))\,a'(x) + \int_{a(x)}^{b(x)} \frac{\partial}{\partial x} f(x,t)\,dt.$$

$$\Rightarrow \frac{\partial}{\partial b_i}\int_a^b (f-f_n)H(x-b_j)\,dx = \frac{\partial}{\partial b_i}\int_{b_j}^b (f-f_n)\,dx$$

$$= \begin{cases} \int_{b_j}^b \frac{\partial}{\partial b_i}(f-f_n)\,dx, & i\ne j \\ (f_n-f)(b_j)+\int_{b_j}^b \frac{\partial}{\partial b_i}(f-f_n)\,dx, & i=j\end{cases} = \begin{cases}\int_{b_j}^b c_i H(x-b_i), = c_i\int_a^b H_iH_j, & i\ne j\\ (f_n-f)(b_j)+c_i\int_a^b H_iH_j\,dx, & i=j.\end{cases}$$

Iterative Method for

$$\begin{cases} 0=\nabla_{\vec c}\, J(f_n) = M(\vec b)\,\vec c - \vec F(\vec b)\\ 0 = \nabla_{\vec b}\, J(f_n) = D(\vec c)\int_a^b (f-f_n)\vec H\,dx\end{cases}$$

$\vec b=(b_1,\cdots,b_n)$ — mesh points

$\vec c=(c_0,c_1,\cdots,c_n)$ — coefficients

- initialization $\vec b^{(0)}$: $\quad b_i^{(0)} = a + i\,\frac{b-a}{n+1}$ for $i=0,1,\cdots,n$

for $k=1,2,\cdots$

- linear parameter $\vec c^{(k)}$: $\quad \vec c^{(k)} = M^{-1}(\vec b^{(k-1)})\,\vec F(\vec b^{(k-1)})$ $\boxed{\cancel{\text{for } k=1,2,\cdots}}$

- nonlinear parameter $\vec b^{(k)}$ (Newton / Gauss-Newton)

$$\vec b^{(k)} = \vec b^{(k-1)} - \omega_k\, \nabla^2_{\vec b} J\left(f_n(\cdot;\vec c^{(k)},\vec b^{(k-1)})\right)\nabla_{\vec b} J\left(f_n(\cdot;\vec c^{(k)},\vec b^{(k-1)})\right)$$

$$= \vec b^{(k-1)} - \omega_k \left[D(\vec w^{(k-1)}) + A(\vec b^{(k-1)})\,D(\vec c^{(k)})\right]^{-1}\int_a^b \left(f-f_n(x;\vec c^{(k)},\vec b^{(k-1)})\right)\vec H(x;\vec b^{(k-1)})\,dx.$$

Lemma $A(\vec b)$ is symmetric positive definite.