Regular Families of Kernels for Nonlinear Approximation

Keaton Hamm

Vanderbilt University

Southeastern Analysis Meeting, 2017

Basics

- Joint work with Jeff Ledford (Virginia Commonwealth University)

Basics

- Joint work with Jeff Ledford (Virginia Commonwealth University)

Definition

The L_p *approximation error* of approximating a function f from a space X is given by

$$\mathcal{E}(f,X)_{L_p} := \inf_{g\in X} \|f-g\|_{L_p}.$$

Basics

- Joint work with Jeff Ledford (Virginia Commonwealth University)

Definition

The L_p *approximation error* of approximating a function f from a space X is given by

$$\mathcal{E}(f, X)_{L_p} := \inf_{g \in X} \|f - g\|_{L_p}.$$

Definition

The *Triebel–Lizorkin space* of functions with smoothness $s \in \mathbb{R}_+$ can be defined as follows: let $\phi \in \mathscr{S}(\mathbb{R}^d)$ have $\lfloor s \rfloor$ vanishing moments and $\phi_j(x) := 2^{jd}\phi(2^j x)$, then

$$F^s_{\tau,q} := \left\{ f \in \mathscr{S}'(\mathbb{R}^d) : \|f\|_{F^s_{\tau,q}} := \left\| \left(\sum_{j=0}^{\infty} |2^{sj}\phi_j * f|^q \right)^{\frac{1}{q}} \right\|_{L_\tau(\mathbb{R}^d)} < \infty \right\}.$$

Basics

Note:

(1) The Definition of $F^s_{\tau,q}$ does not depend on ϕ.

Basics

Note:

(1) The Definition of $F^s_{\tau,q}$ does not depend on ϕ.

(2) Every $f \in F^s_{\tau,q}$ has a wavelet expansion $f = \sum_{I \in \mathcal{D}} f_I \phi_I$ ($\mathcal{D} = \{$dyadic cubes$\}$).

Note:

(1) The Definition of $F^s_{\tau,q}$ does not depend on ϕ.

(2) Every $f \in F^s_{\tau,q}$ has a wavelet expansion $f = \sum_{I\in\mathcal{D}} f_I\phi_I$ ($\mathcal{D} = \{$dyadic cubes$\}$).

(3) Equivalently, $F^s_{\tau,q} := \{f : \|f\|_{F^s_{\tau,q}} := \|M_{s,q}f\|_{L_\tau(\mathbb{R}^d)} < \infty\}$, with

$$M_{s,q}f(t) := \left(\sum_{I\in\mathcal{D}} |I|^{-\frac{sq}{d}}|f_I|^q\chi_I(t)\right)^{\frac{1}{q}}.$$

History

Theorem (DeVore, Jawerth, Popov, '92)

Let ϕ *define* $F^s_{\tau,q}$*, and let*

$$\Sigma_N := \left\{ \sum_{(j,k)\in\Lambda} c_{j,k}\phi(2^k\cdot -j) : |\Lambda| = N, (c_{j,k}) \subset \mathbb{C} \right\}.$$

Let $p \in (0,\infty)$*. If* $\tau = (\frac{1}{p} - \frac{s}{d})^{-1}$ *and* $q = (1+\frac{s}{d})^{-1}$*, then for every* $f \in F^s_{\tau,q}$*,*

$$\mathcal{E}(f,\Sigma_N)_{L_p} \leq CN^{-\frac{s}{d}}\|f\|_{F^s_{\tau,q}}.$$

History

Theorem (DeVore, Ron, '10)

For certain ϕ *(e.g.* surface splines *defined by* $\phi_m(x) = |x|^{2m-d}$ *if* d *is odd and* $\phi_m(x) = |x|^{2m-d} \ln |x|$ *if* d *is even) let*

$$\Sigma_N(\phi) := \left\{ \sum_{j=1}^{N} a_j \phi(\cdot - j) : (a_j) \subset \mathbb{C},\ (x_j) \subset \mathbb{R}^d \right\}.$$

Let $p \in [1, \infty)$ *and* τ *and* q *as before. Then for every* $f \in F^s_{\tau,q}$,

$$\mathcal{E}(f, \Sigma_N(\phi))_{L_p} \leq C N^{-\frac{s}{d}} \|f\|_{F^s_{\tau,q}}.$$

History

Theorem (Hangelbroek, Ron '10)

Let

$$\mathbb{G}_N := \left\{ \sum_{j=1}^{N} a_j exp\left(-\left| \frac{\cdot - x_j}{\sigma_j} \right|^2 \right) : (a_j) \subset \mathbb{C},\ (x_j) \subset \mathbb{R}^d,\ (\sigma_j) \subset \mathbb{R}_+ \right\},$$

and $p \in [1, \infty)$*. Then for* τ, q *as above, every* $f \in F^s_{\tau,q}$ *satisfies*

$$\mathcal{E}(f, \mathbb{G}_N)_{L_p} \leq CN^{-\frac{s}{d}} \|f\|_{F^s_{\tau,q}}.$$

Problem Statement

Problem

What are conditions on families $(\phi_\alpha)_{\alpha\in A}$, $A\subset\mathbb{R}_+$ *such that the approximation space*

$$\Phi_N := \left\{\sum_{j=1}^N a_j\phi_{\alpha_j}(\cdot - x_j) : (a_j)\subset\mathbb{C},\ (x_j)\subset\mathbb{R}^d,\ (\alpha_j)\subset A\right\}$$

is such that $\mathcal{E}(f,\Phi_N)_{L_p}\leq CN^{-\frac{s}{d}}\|f\|_{F^s_{\tau,q}}$ *for the appropriate parameters* p, s, τ, *and* q*?*

Decaying Kernel Conditions

(A1) Φ_N is closed under translations and dilations

(A2) ϕ_α is continuous, and $\sup\limits_{\alpha\in A}\|\phi_\alpha\|_{L_\infty} \leq C$.

(A3) $\forall k \in \mathbb{N}\cup\{0\}$ and for sufficiently large $\alpha(=\alpha(k))$,

$$\sum_{j\neq 0}\left\|\frac{\widehat{\phi_\alpha}(\cdot + h^{-1}j)}{\widehat{\phi_\alpha}(\cdot)}\right\|_{L_\infty(B)} \leq Ch^k,$$

where C is independent of α.

(A4) $\forall k \in \mathbb{N}$ and for sufficiently large α, $D^\gamma(1/\widehat{\phi_\alpha}) \in L_2(B)$, for all $|\gamma| \leq k$.

(A5) $\forall k \in \mathbb{N}$ and for sufficiently large α, $|\phi_\alpha(x)| = O(|x|^{-2k}),\ |x| \to \infty$.

(A6) The Poisson Summation Formula holds for ϕ_α for every $\alpha \in A$.

Decaying Kernel Conditions

Theorem ((H, Ledford))

*If $(\phi_\alpha)_{\alpha\in A}$ satisfies **(A1)–(A6)**, then if $p \in [1,\infty]$, $\tau = (\frac{1}{p} - \frac{s}{d})^{-1}$ and $q = (1 + \frac{s}{d})^{-1}$,*

$$\mathcal{E}(f, \Phi_N)_{L_p} \leq CN^{-\frac{s}{d}} \|f\|_{F^s_{\tau,q}}, \quad \forall f \in F^s_{\tau,q}.$$

Examples

Gaussians: $\phi_\alpha(x) := e^{-\left|\frac{x}{\alpha}\right|^2}, \quad \alpha \in \mathbb{R}_+$

Examples

Gaussians: $\phi_\alpha(x) := e^{-\left|\frac{x}{\alpha}\right|^2}, \quad \alpha \in \mathbb{R}_+$

Inverse Multiquadrics:
$\phi_{\alpha,c}(x) := (|x|^2 + c^2)^{-\alpha},\ \alpha \in (d + 1/2, \infty),\ c \in \mathbb{R}_+$

Examples

Gaussians: $\phi_\alpha(x) := e^{-\left|\frac{x}{\alpha}\right|^2}, \quad \alpha \in \mathbb{R}_+$

Inverse Multiquadrics:
$\phi_{\alpha,c}(x) := (|x|^2 + c^2)^{-\alpha},\ \alpha \in (d + 1/2, \infty),\ c \in \mathbb{R}_+$

Matérn Kernels: $\rho_{\alpha,c}(x) := \widehat{\phi}_{\alpha,c}(x) = C|x|^{-\alpha-\frac{d}{2}} K_{\alpha-\frac{d}{2}}(c|x|)$

Examples

Gaussians: $\phi_\alpha(x) := e^{-\left|\frac{x}{\alpha}\right|^2}, \quad \alpha \in \mathbb{R}_+$

Inverse Multiquadrics:
$\phi_{\alpha,c}(x) := (|x|^2 + c^2)^{-\alpha},\ \alpha \in (d + 1/2, \infty),\ c \in \mathbb{R}_+$

Matérn Kernels: $\rho_{\alpha,c}(x) := \widehat{\phi}_{\alpha,c}(x) = C|x|^{-\alpha-\frac{d}{2}} K_{\alpha-\frac{d}{2}}(c|x|)$

Cardinal Functions: Given ϕ_α, define $\widehat{L}_\alpha(\xi) := \frac{\widehat{\phi_\alpha}(\xi)}{\sum_{j\in\mathbb{Z}} \widehat{\phi_\alpha}(\xi-j)}$

Growing Kernels

Main Idea: Use divided differences to impose decay.

Growing Kernels

Main Idea: Use divided differences to impose decay.

Examples:

Multiquadrics: ϕ_α, the 2α-th divided difference of $(x^2+\alpha^2)^{\alpha-\frac{1}{2}}$

Growing Kernels

Main Idea: Use divided differences to impose decay.

Examples:

Multiquadrics: ϕ_α, the 2α-th divided difference of $(x^2 + \alpha^2)^{\alpha - \frac{1}{2}}$

Power Functions: ϕ_α, the 2α-th divided difference of $|x|^\alpha$.

Numerical Experiments

Take ψ to be the Meyer wavelet (infinitely differentiable, bandlimited):

$$\widehat{\psi}(\xi) := \begin{cases} \sin\left(\frac{\pi}{2}\nu\left(\frac{3|\xi|}{2\pi}-1\right)\right)e^{i\frac{\xi}{2}}, & \frac{2\pi}{3}\leq|\xi|\leq\frac{4\pi}{3}, \\ \cos\left(\frac{\pi}{2}\nu\left(\frac{3|\xi|}{4\pi}-1\right)\right)e^{i\frac{\xi}{2}}, & \frac{4\pi}{3}\leq|\xi|\leq\frac{8\pi}{3}, \\ 0, & \text{otherwise}, \end{cases}$$

Numerical Experiments

Take ψ to be the Meyer wavelet (infinitely differentiable, bandlimited):

$$\widehat{\psi}(\xi) := \begin{cases} \sin\left(\frac{\pi}{2}\nu\left(\frac{3|\xi|}{2\pi}-1\right)\right)e^{i\frac{\xi}{2}}, & \frac{2\pi}{3}\leq|\xi|\leq\frac{4\pi}{3}, \\ \cos\left(\frac{\pi}{2}\nu\left(\frac{3|\xi|}{4\pi}-1\right)\right)e^{i\frac{\xi}{2}}, & \frac{4\pi}{3}\leq|\xi|\leq\frac{8\pi}{3}, \\ 0, & \text{otherwise}, \end{cases}$$

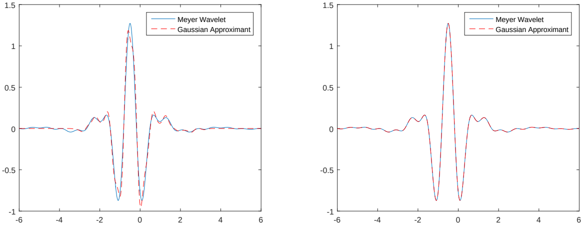

Figure: Gaussian approximation of Meyer mother wavelet with budget $N = 10$ (left) and $N = 50$ (right)

Numerical Experiments

Random 5–term wavelet expansion from the Meyer wavelet:

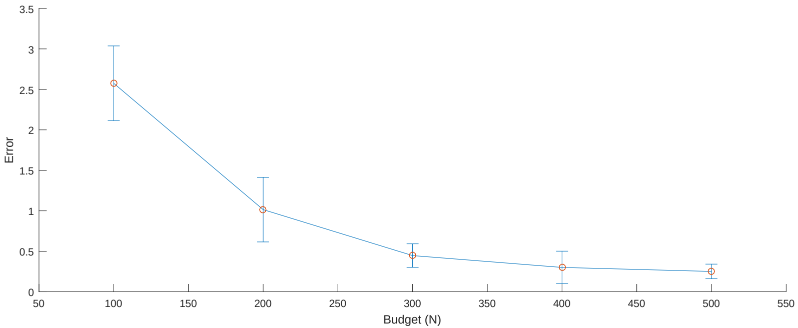

Figure: Average L_∞ error vs. N for Gaussian approximation with $\alpha = 0.2$ of random 5–term wavelet

Drawbacks/ Future Work

(1) The approximant giving the theorems above require knowledge of wavelet coefficients $f = \sum_{I \in \mathcal{D}} f_I \psi_I$.

Drawbacks/ Future Work

(1) The approximant giving the theorems above require knowledge of wavelet coefficients $f = \sum_{I\in\mathcal{D}} f_I\psi_I$.

(2) Making approximant is computationally inefficient.

Drawbacks/ Future Work

(1) The approximant giving the theorems above require knowledge of wavelet coefficients $f = \sum_{I \in \mathcal{D}} f_I \psi_I$.

(2) Making approximant is computationally inefficient.

(3) Question: Is there a computationally feasible method to make $S_{f,N} \in \Phi_N$ achieving these approximation rates of $O(N^{-s/d})$?

Thanks!