Kernel methods and parametric PDEs

Christian Rieger

partly based on joint work with

Michael Griebel (Bonn University) & Barbara Zwicknagl (Würzburg University)

Institute for Numerical Simulation
University of Bonn

supported by DFG through SFB 1060

International Conference of Kernel-Based Approximation Methods in Machine Learning

Outline

Classical reconstruction problem

- Given: data $y_n = f(x_n) + \varepsilon_n$ for $n = 1, \ldots, N$
- centers $X_N := \{x_1, \ldots, x_N\} \subset \Omega \subset \mathbb{R}^d$
- Wanted: approximating function $s_f : \Omega \to \mathbb{R}$

Ansatz:

- $f \in \mathcal{H}(\Omega)$ function space
- typical solution $s_f = \arg\min_{s \in \mathcal{H}(\Omega)} J_{X_N,y}(s)$ with

$$J_{X_N,y}(s) := \sum_{n=1}^{N} (s(x_n) - y_n)^2 + \lambda \|s\|^2_{\mathcal{H}(\Omega)} + \ldots$$

- Useful properties: stability and consistency

$$\|f - s_f\|_{\mathcal{H}} \le \|f\|_{\mathcal{H}} + \frac{1}{\sqrt{\lambda}}(J_{X_N,y}(s_f))^{1/2} \le \|f\|_{\mathcal{H}} + \frac{1}{\sqrt{\lambda}}(J_{X_N,y}(f))^{1/2} \le \frac{1}{\sqrt{\lambda}}\|\varepsilon\|_{\ell_2} + 2\|f\|_{\mathcal{H}}$$

$$\|(f - s_f)|_X\|_{\ell_\infty(X)} \le \|\varepsilon\|_{\ell_\infty} + (J_{X_N,y}(s_f))^{1/2} \le 2\|\varepsilon\|_{\ell_2} + \sqrt{\lambda}\|f\|_{\mathcal{H}}.$$

Reproducing kernel Hilbert space (RKHS)

Definition

Hilbert space $\mathcal{H}(\Omega) \subset C(\Omega)$ is RKHS, if there is $K : \Omega \times \Omega \to \mathbb{R}$, s.t.

- $K(\cdot, x) \in \mathcal{H}(\Omega)$ for all $x \in \Omega$
- $f(x) = (f, K(\cdot, x))_{\mathcal{H}(\Omega)}$ for all $f \in \mathcal{H}(\Omega)$, $x \in \Omega$

Basic properties

- K is symmetric, i.e., for all $x, y \in \Omega$

$$K(x, y) = (K(\cdot, y), K(\cdot, x))_{\mathcal{H}} = (K(\cdot, x), K(\cdot, y))_{\mathcal{H}} = K(y, x)$$

- K is positive semi-definite, i.e., for all $N \in \mathbb{N}$, all $X_N := \{x_1, \ldots, x_N\} \subset \Omega$ and all $c \in \mathbb{R}^N$

$$\begin{aligned}\sum_{i,j=1}^{N} c_i c_j K(x_i, x_j) &= \sum_{i,j=1}^{N} c_i c_j (K(\cdot, x_i), K(\cdot, x_j))_{\mathcal{H}} \\ &= \left\| \sum_{i=1}^{N} c_i K(\cdot, x_i) \right\|_{\mathcal{H}}^2 \geq 0\end{aligned}$$

Optimal reconstruction in an RKHS

Representer Theorem

Let $\lambda > 0$, $X_N = \{x_1, \ldots, x_N\} \subset \Omega$, and $y \in \mathbb{R}^N$. Then a minimizer s^* for

$$J_{X_N,y}(s) := \sum_{j=1}^{N} \left(s\left(x_j\right) - y_j\right)^2 + \lambda \left\|s\right\|^2_{\mathcal{H}(\Omega)}$$

lies in span$\{K(x_1,\cdot), \ldots, K(\cdot, x_N)\}$, i.e., $s^*(x) = \sum_{j=1}^N \alpha_j K(\cdot, x_j)$.

Sketch of the proof: Let $s \in \mathcal{H}$. Decompose $s = s^{\|} + s^{\perp}$ with $s^{\|} \in$ span$\{K(x_1,\cdot), \ldots, K(\cdot, x_N)\}$ and $s^{\|} \perp s^{\perp}$. Then

$$\begin{aligned} J_{X_N,y}(s) &:= \sum_{j=1}^{N} \left(s\left(x_j\right) - y_j\right)^2 + \lambda \|s\|^2_{\mathcal{H}(\Omega)} \\ &= \sum_{j=1}^{N} \left(\left(s^{\|}, K(x_j,\cdot)\right)_{\mathcal{H}} - y_j\right)^2 + \lambda \left(\left\|s^{\|}\right\|^2_{\mathcal{H}(\Omega)} + \left\|s^{\perp}\right\|^2_{\mathcal{H}(\Omega)}\right) \end{aligned}$$

(Schölkopf and Smola 2002)

Characteristics of Kernel-Based Approximation

Advantages

- Optimality properties („splines“)
- Scattered data
- Generalized recovery (not only point evaluations) via Riesz representation

Problems

- Optimality holds only in associated Hilbert space
- Problem-induced kernel often not available in closed form
- Numerically feasible approximation necessary
- Often interested in multiscale decompositions

Parametric PDE

Parametric Poisson Problem

$$
\begin{aligned}
-\operatorname{div}\left(a\left(y,x\right)\nabla u\left(y,x\right)\right) = g\left(x\right) \quad &\text{for all } \left(y,x\right)\in\Omega\times D\\
u\left(y,x\right) = 0 \quad &\text{for all } \left(y,x\right)\in\Omega\times\partial D
\end{aligned}
$$

Quantity of interest: $f:\Omega\to\mathbb{R}$, $f(y)=\langle Q,u(y,\cdot)\rangle_{V^\star(D),V(D)}$

- $Y_{N_S}=\{y_1,\ldots,y_{N_S}\}\subset\Omega\subset\mathbb{R}^{N_P}$ with N_P large
- a such that pde is well posed
- data: $(y_k,z_k)\in\Omega\times\mathbb{R}$, with $f(y_k)\approx z_k$, $1\le k\le N_S$
- exploit structure in reconstruction of f
- determine kernel from problem data, i.e. $a\in L^\infty\left(\Omega,C^1(\overline{D})\right)$

Point evaluation

- We need to evaluate $f(y_k)$ numerically

Discretization in space

- trial space $V_h := \text{span}\{\Psi_1, \ldots, \Psi_{N_h}\} \subset H_0^1(D)$
- find $u^h(y_k, \cdot)$ such that

$$\int_D a(y_k, x) \nabla u^h(y_k, x) \nabla v_h(x)\, dx = \int_D f(x) v_h(x)\, dx$$

for all $v_h \in V_h$.

- $\epsilon_h := \max_{k=1}^{N_S} \left| \langle Q, u^h(y_k, \cdot) - u(y_k, \cdot) \rangle_{H^{-1}(D), H_0^1(D)} \right|$
- ϵ_h is a parameter and we can achieve $\epsilon_h \to 0$

see Mike's talk yesterday

- high (N_P) dimensional reconstruction problem for $f : \Omega \to \mathbb{R}$ from (y_k, z_k), $1 \leq k \leq N_S$
- deterministically polluted data (ϵ_h)
- we study reconstruction of f, not quadrature $\to$ ill-conditioned problem
- need to discuss N_S as function of N_P
- 1st ingredient: smoothness of $f \to$ RKHS
- 2rd ingredient: error estimates which allow (a priori) coupling of discretization parameters

Parametric regularity

Separation of variables:

$$u(y,x) = \sum_{k=1}^{\infty} \sum_{\substack{\nu \in \mathbb{N}^{N_P} \\ |\nu|=k}} u_\nu(x)\, \Phi_\nu(y)$$

Fast decaying coefficients w.r.t. function space $V(D) := H_0^1(D)$

Cohen, DeVore, Schwab (2010), Babuška,Tempone, Zouraris (2004), Babuška, Nobile, Tempone (2007), Dung, Griebel, Huy, CR (2017) . . .

$\|u_\nu\|_{V(D)} \le C(g,\delta)\rho^{-\nu}$ for a δ admissible vector $\rho \in \mathbb{R}^{N_P}$

Taylor

- $u_\nu(x) := \frac{1}{\nu!}\partial_y^\nu u(0,x)$
- $\Phi_\nu(y) := y^\nu$ Monomials

Power series kernels
Zwicknagl (2009), Zwicknagl &Schaback (2013)

Spectral

- $u_\nu(x) := \int_D u(y,x)\,\Phi_\nu(y)\,dy$
- $\Phi_\nu(\cdot)$ orthogonal system w.r.t. to the joint probability density

Mercer kernels

RKHS containing f

- $f(y) = \langle Q, u(y,\cdot)\rangle_{V^\star(D),V(D)}$
- $u(y,x) = \sum_{k=1}^{\infty} \sum_{\substack{\nu\in\mathbb{N}^{N_P} \\ |\nu|=k}} u_\nu(x)\,\Phi_\nu(y)$

leads to

$$f(y) = \sum_{k=1}^{\infty} \sum_{\substack{\nu\in\mathbb{N}^{N_P} \\ |\nu|=k}} f_\nu \Phi_\nu(y) \quad \text{with} \quad f_\nu = \langle Q, u_\nu\rangle_{V^\star(D),V(D)}$$

and bounds

$$|f_\nu| \leq \|Q\|\,\|u_\nu\| \leq C(g,\delta)\,\|Q\|\,\rho^{-\nu}$$

f is contained in an RKHS of a power series kernel

$$K(y,\tilde{y}) = \sum_{\nu\in\mathbb{N}^{N_P}} \lambda_\nu \tilde{y}^\nu y^\nu \quad \text{sometimes closed formula}$$

Griebel&CR (2017) Hardy-spaces as RKHS

Reproduction formula = Taylor expansion

The weighted ℓ_2-space

$$\mathcal{H} := \left\{ f(\cdot) = \sum_{\alpha \in \mathbb{N}_0^{N_P}} a_\alpha (\cdot)^\alpha \text{ with } \sum_{\alpha \in \mathbb{N}_0^{N_P}} \frac{\alpha!^2}{w_\alpha} a_\alpha^2 < \infty \right\}$$

with inner product

$$(f,g)_{\mathcal{H}} := \sum_{\beta \in \mathbb{N}_0^{N_P}} \frac{1}{w_\beta} (D^\beta f(0))(D^\beta g(0))$$

is the native space for $K(x,y) := \sum_{\alpha \in \mathbb{N}_0^{N_P}} w_\alpha \frac{x^\alpha y^\alpha}{\alpha!^2}$.

$$\begin{aligned}
(f, K(\cdot, y))_{\mathcal{H}} &= \sum_{\alpha \in \mathbb{N}_0^{N_P}} \frac{(D^\alpha f(0))\left(D_1^\alpha K(0,y)\right)}{w_\alpha} = \sum_{\alpha \in \mathbb{N}_0^{N_P}} \frac{1}{w_\alpha} (a_\alpha \alpha!) \left(w_\alpha \frac{y^\alpha}{\alpha!} \right) \\
&= \sum_{\alpha \in \mathbb{N}_0^{N_P}} a_\alpha y^\alpha = f(y) .
\end{aligned}$$

Zwicknagl (2009), Zwicknagl &Schaback (2013)

Compuation of kernels

Example: $H^1_\gamma([a,b])$

$$(f,g)_{H^1_\gamma([a,b])} = \int_a^b f(x)\, g(x)\, dx + \gamma \int_a^b \partial f(x) \partial g(x) dx$$

similar to anchored spaces $H_{1,\gamma}$ from Dick, Kuo, Sloan (2014)

Aim: reproduction formula

$$\begin{aligned} f(y) &= (f, K(\cdot, y))_{H^1_\gamma([a,b])} \\ &= \int_a^b f(x)\, K(x,y)\, dx + \gamma \int_a^b \partial f(x) \partial_x K(x,y)\, dx \end{aligned}$$

Fasshauer, Ye (2011), Cavoretto, Fasshauer, McCourt (2017)

Greens function

- integration by parts

$$\begin{aligned}(f, g)_{H^1_\gamma([a,b])} &= \int_a^b f(x)\left(g(x) - \gamma\partial^2 g(x)\right) dx \\ &+ \gamma\left[f(b)\partial g(b) - f(a)\partial g(a)\right]\end{aligned}$$

- Greens function for

$$\mathcal{D} = \mathrm{Id} - \gamma\partial^2$$

  subject to Neumann BC

- Eigensystem of Neumann Laplace

$$\phi_k(x) = \cos\left(\pi k \frac{x-a}{b-a}\right) \quad \lambda_k = (\pi k)^2$$

- Mercer-kernel

$$K(x, y) = \sum_{k \in \mathbb{N}} \frac{1}{1 + \gamma(\pi k)^2} \cos\left(\pi k \frac{x-a}{b-a}\right) \cos\left(\pi k \frac{y-a}{b-a}\right)$$

Explicit computations

assume: ∂K diagonal jump

$$\begin{aligned}
f(x) &= \int_a^b f(y)\,K(x,y)\,dy + \gamma \int_a^b \partial f(y) \partial K(x,y)\,dy \\
&= \int_a^b f(y)\,K(x,y)\,dy - \gamma \int_a^b f(y) \partial^2 K(y,x)\,dy \\
&\quad +\gamma \left[f(y)\partial K(x,y)\right]_a^x + \gamma \left[f(y)\partial K(x,y)\right]_x^b \\
&= \int_a^b f(y)\left(K(x,y) - \gamma \partial^2 K(y,x)\right) dy \\
&\quad \gamma f(y)\left(\partial K(x_+,x) - \partial K(x_-,x)\right) \\
&\quad +\gamma\left(f(b)\partial K(b,x) - f(a)\partial K(a,x)\right.
\end{aligned}$$

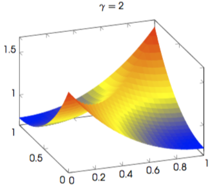
Schaback's Lecture Notes (2007)

COMPUTATION CONT.

$$\begin{aligned} \text{for } a &\leq y \leq x \\ K(y,x) &= \gamma\partial^2 K(y,x) \\ 0 &= \partial K(a,x) \\ \alpha(x) &= \partial K(x_-,x) \end{aligned} \qquad \begin{aligned} \text{for } x &\leq y \leq b \\ K(y,x) &= \gamma\partial^2 K(y,x) \\ 0 &= \partial K(b,x) \\ \beta(x) &= \partial K(x_+,x) \end{aligned}$$

JUMP CONDITION

$$\alpha(x) - \beta(x) = \frac{1}{\gamma}$$



SOLUTION FOR $[0,1]$

$$K(x,y) = \frac{\sqrt{\gamma}}{\sinh\sqrt{\gamma}} \cosh\left(\sqrt{\gamma}\,(1 - \max\{x,y\})\right) \cosh\left(\sqrt{\gamma}\min\{x,y\}\right)$$

Schaback's Lecture Notes (2007)

Manifold Regularization

$$
\begin{aligned}
s^\star &= \mathrm{argmin}_{s\in\mathcal{H}} F(s) \\
F(s) &= \sum_{i=1}^{N_S} V(y_i, z_i, s(y_i)) + \gamma_K \|s\|^2_{\mathcal{H}} + \gamma_{\mathcal{M}} \|s\|^2_{\mathcal{M}}
\end{aligned}
$$

- V loss function and $\mathcal{H}$ RKHS with kernel K
- $\|s\|^2_{\mathcal{M}} = (s, (-\Delta_{\mathcal{M}})s)_{L^2} \to$ Graph-Laplacian $\Delta_{\mathcal{M}}(h)$

$$
\frac{s|^T_{Y_{N_S}} \Delta_{\mathcal{M}}(h) s|_{Y_{N_S}}}{N_S^2} = \frac{1}{N_S^2} \sum_{i,j=1}^{N_S} (s(y_i) - s(y_j))^2 W_{i,j}
$$

- Representer Theorem (Belkin, Niyogi, Sindhwani (2006))

$$
s^\star = \sum_{i=1}^{N_S} \alpha_i K(\cdot, y_i)
$$

Graph Laplace

Similarity measure: $W_{i,j} = \exp\left(-\frac{\|y_i - y_j\|_2}{4t}\right)$

$$\Delta_{\mathcal{M}}(h) = D - W = (L_{j,k})_{1 \le j,k \le N_S}$$

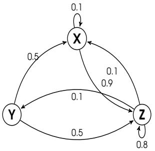
$$L_{j,k} = \begin{cases} \exp\left(-\frac{\|y_j - y_k\|_2}{4t}\right) & \text{if } j \neq k \\ -\sum_{\substack{\ell=1 \\ \ell \neq k}}^{N_S} \exp\left(-\frac{\|y_\ell - y_k\|_2}{4t}\right) & \text{if } j = k \end{cases}$$

- discretization of heat flow
- parameter t has to be chosen carefully
- heat kernel

Heat-kernel induced geometries

- Anistropic heat equation

$$\partial_t u(t,x) - \mathcal{L}_x u(t,x) = 0, \quad u(0,x) = u^{(0)}(x)$$



- $\mathcal{L}_x$ second order elliptic operator,
  e.g. $\mathcal{L}_x u(t,x) = \mathrm{Div}(a(x)\nabla u(t,x))$
- $p(t,x,y)$ fundamental solution

$$\partial_t p(t,x,y) - \mathcal{L}_x p(t,x,y) = 0, \qquad \lim_{t\to 0} p(t,x,y) = \delta(x-y)$$

- Diffusion distance

$$\begin{aligned} D_t^2(x_1,x_2) &:= \|p(t,x_1,\cdot) - p(t,x_2,\cdot)\|^2_{L^2(d\mu)} \\ &= \sum_{\ell=0}^{\infty} e^{-2\mu_\ell t}(\phi_\ell(x_1) - \phi_\ell(x_2))^2 \end{aligned}$$

  using eigenvalues and eigenfunctions of $\mathcal{L}_x$
- Kernels

$$K(x,y) = \sum_{\ell\in\mathbb{N}} \lambda_\ell \phi_\ell(x)\phi_\ell(y)$$

Coifman & Lafon ('06)

Typical questions for positive definite multiscale kernels

$$K(x, y) = \sum_{\ell \in \mathbb{N}} \lambda_\ell \phi_\ell(x) \phi_\ell(y)$$

Want to reconstruct

$$f \in \mathcal{H} := \{f = \sum_{\ell=0}^{\infty} a_\ell \lambda_\ell^{1/2} \phi_\ell : \sum_{\ell=0}^{\infty} |a_\ell|^2 < \infty\}$$

from values $y_n \approx f(x_n)$, $n = 1, \ldots, N$

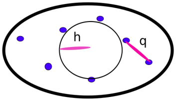
Ansatz: $Rf(x) = \sum_{n=1}^{N} f_n K(x, x_n)$

- Bound the worst case approximation error for a function $f \in \mathcal{H}$ in dependence on the points x_n
- Approximate K by finite sums $K^L(x, y) = \sum_{\ell=0}^{L} \lambda_\ell \phi_\ell(x) \phi_\ell(y)$ such that
  - the Gramian matrix $(K^L(x_n, x_{\tilde{n}}))_{n, \tilde{n} \in \mathbb{N}}$ remains invertible.
  - the worst-case error for an arbitrary admissible (price) function has the same asymptotic behavior when using the truncated kernel

SOME RESULTS

$$K(x,y) = \sum_{\ell=0}^{L} \lambda_\ell \phi_\ell(x)\phi_\ell(y)$$

- Sufficient conditions for L, e.g., if the weights λ_ℓ decay geometrically, then choosing $L = L(q_X)$ preserves the required good properties.
  Griebel, Rieger & BZ ('15) & ('17)



- Error analysis for generalized Besov-spaces on metric measure spaces, stability results Griebel, CR & Zwicknagl ('15) & ('17)
- If $\{\phi_\ell\}$ is not an orthogonal system but e.g. monomials, tight frames, ...
  Zwicknagl ('09); Schaback & Zwicknagl ('13); Opfer ('05); Griebel, CR & Zwicknagl ('15) & ('17)

Deterministic Error Analysis

$$
\begin{aligned}
s^\star &= \mathrm{argmin}_{s\in\mathcal{H}} F(s) \\
F(s) &= \sum_{i=1}^{N_S} V(y_i, z_i, s(y_i)) + \gamma_K \|s\|_{\mathcal{H}}^2 + \gamma_{\mathcal{M}} \|s\|_{\mathcal{M}}^2
\end{aligned}
$$

- Stability

$$
\|s^\star\|_{\mathcal{H}}^2 \le \|f\|_{\mathcal{H}}^2 + \frac{\gamma_{\mathcal{M}}}{\gamma_K}\|f\|_{\mathcal{M}}^2 + \frac{1}{\gamma_K}\sum_{i=1}^{N_S} V(y_i, z_i, f(y_i))
$$

- Consistency

$$
V(y_i, z_i, s^\star(y_i)) \le \gamma_K \|f\|_{\mathcal{H}}^2 + \gamma_{\mathcal{M}} \|f\|_{\mathcal{M}}^2 + \sum_{i=1}^{N_S} V(y_i, z_i, f(y_i))
$$

- Stochastic Analysis in L_∞, Fischer, Steinwart (2017)

- least-squares $V(y_i, z_i, s(y_i)) = (z_i - s(y_i))^2$
- $f \in \mathcal{H}$ and $|z_i - f(y_i)| \leq \epsilon_h$
- Stability

$$\|s^\star\|_{\mathcal{H}}^2 \leq \|f\|_{\mathcal{H}}^2 + \frac{\gamma_{\mathcal{M}}}{\gamma_K}\|f\|_{\mathcal{M}}^2 + \frac{N_S \epsilon_h^2}{\gamma_K}$$

- Consistency

$$(f(y_i) - s^\star(y_i))^2 \leq \gamma_K \|f\|_{\mathcal{H}}^2 + \gamma_{\mathcal{M}}\|f\|_{\mathcal{M}}^2 + N_S \epsilon_h^2$$

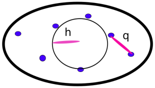
Sampling inequalities for scattered data

$$\|u\|_W \lesssim C_1\left(h_{Y_{N_S}}\right)\|u\|_S + C_2\left(h_{Y_{N_S}}\right)\left\|u|_{h_{Y_{N_S}}}\right\|_{W(Y_{N_S})}$$

$C_1 \to 0$ as $h_{Y_{N_S}} \to 0$, smoothness difference W and S

- Narcowich, J.D. Ward & Wendland (2005)
  functions with zeros, $u|_{Y_{N_S}} = 0$,

  $$|u|_{W_p^k} \leq Ch^{m-k}|u|_{W_p^m}$$



- Wendland & CR (2005)

  $$|u|_{W_p^k} \leq C\left(h^{m-k}|u|_{W_p^m} + h^{-k}\left\|u|_{Y_{N_S}}\right\|_{\ell_\infty(Y_{N_S})}\right)$$

- Madych, J. Approx. Th. (2006)

  $$\|u\|_{L_p(\Omega)} \leq C\left(h^m|u|_{W_p^m(\Omega)} + h^{d/p}\|u|_X\|_{\ell_p(Y_{N_S})}\right)$$

- and many more ...

Exponential Approximation Rates

Smooth functions

Hilbert space $\mathcal{H}$ of functions $f : \Omega \to \mathbb{R}$ with embedding constants $C_{\mathcal{H}}(k)$, s. t. for a fixed p and all $f \in \mathcal{H}$

$$\|f\|_{W_p^k(\Omega)} \leq C_{\mathcal{H}}(k)\|f\|_{\mathcal{H}} .$$

The sampling rate

1. is measured in $h = h_{Y_{N_S},\Omega} = \sup_{y\in\Omega} \min_{y_j \in Y_{N_S}} \|y - y_j\|$
2. depends on the embedding constants $C_{\mathcal{H}}(k)$ for $k \to \infty$
3. depends on the geometry of the domain Ω

Guideline

The slower the embedding constants $C_{\mathcal{H}}(k)$ grow with $k \to \infty$, the better are the approximation rates!

Exponential Rates: Results for Gaussians

Embedding constants

$C_{\mathcal{H}}(k) \leq C_E^k k^{(1-\frac{1}{2})k}$ for all $k \in \mathbb{N}$.

- $\mathcal{D}$ compact cube CR & Zwicknagl, (2010)

$$\|r\|_{L_q(\mathcal{D})} \leq e^{C\log(h)/h} \|r\|_{\mathcal{H}} + c^{1/h} \|r|_X\|_{\ell_\infty(Y_{N_S})}$$

- $\mathcal{D}$ interior cone condition, points clustered towards boundary CR & Zwicknagl, (2014)

$$\|D^\alpha r\|_{L_q(\mathcal{D})} \leq e^{C\log(h)/h}\|r\|_{\mathcal{H}} + ch^{-2|\alpha|}\|r|_{Y_{N_S}}\|_{\ell_\infty(Y_{N_S})}$$

Interpolation rate

$$\left\|f - s_{f,Y_{N_S}}\right\|_{L_q(\mathcal{D})} \leq e^{C\log(h)/h} \|f\|_{\mathcal{H}}$$

Application to interpolation

- $r = f - s_{f,Y_S}$
- $\left\|s_{f,Y_{N_S}}\right\|_{\mathcal{H}} \leq \|f\|_{\mathcal{H}}$ and $r|_{Y_{N_S}} = 0$

Deterministic error analysis for manifold semi-supervised learning

Ingredients

- $$\|s^\star\|_{\mathcal{H}}^2 \leq \|f\|_{\mathcal{H}}^2 + \frac{\gamma_{\mathcal{M}}}{\gamma_K}\|f\|_{\mathcal{M}}^2 + \frac{N_S \epsilon_h^2}{\gamma_K}$$
- $$f(y_i) - s^\star(y_i))^2 \leq \gamma_K \|f\|_{\mathcal{H}}^2 + \gamma_{\mathcal{M}}\|f\|_{\mathcal{M}}^2 + N_S \epsilon_h^2$$
- $$\|u\|_W \lesssim C_1\left(h_{Y_{N_S}}\right)\|u\|_S + C_2\left(h_{Y_{N_S}}\right)\left\| u|_{h_{Y_{N_S}}} \right\|_{W(Y_{N_S})}$$
- choose $W = L^\infty(\Omega) \Rightarrow C_2\left(h_{Y_{N_S}}\right) = \text{const}$
- choose $u = f - s^\star \in \mathcal{H}$

Griebel&CR (2017)

$$
\begin{aligned}
\|f - s^\star\|_{L^\infty(\Omega)} &\lesssim C_1\left(h_{Y_{N_S}}\right)\|f - s^\star\|_{\mathcal{H}} + C_2\|f|_{Y_{N_S}} - s^\star|_{Y_{N_S}}\|_{\ell^\infty(Y_{N_S})} \\
&\lesssim C_1\left(h_{Y_{N_S}}\right)\left(2 + \sqrt{\frac{\gamma_K}{\gamma_{\mathcal{M}}}}\right)\|f\|_{\mathcal{H}} \\
&+ C_2\left(\sqrt{\gamma_K}\|f\|_{\mathcal{H}} + \sqrt{\gamma_K}\|f\|_{\mathcal{M}}\right) \\
&+ \left(1 + \frac{C_1\left(h_{Y_{N_S}}\right)}{\sqrt{\gamma_K}}\right)\sqrt{N_S}\epsilon_h
\end{aligned}
$$

Choose:

$$
\sqrt{\gamma_K} = \sqrt{\gamma_{\mathcal{M}}} \sim C_1\left(h_{Y_{N_S}}\right)
$$

$$
\|f - s^\star\|_{L^\infty(\Omega)} \lesssim C_1\left(h_{Y_{N_S}}\right)\left(\|f\|_{\mathcal{M}} + \|f\|_{\mathcal{H}}\right) + \sqrt{N_S}\epsilon_h
$$

Griebel&CR (2017)

Conclusion

- problem adapted kernels
- problems from parametric pde, greens function, heat kernels
- kernel needs careful computation
- determinitsic errors need regularization
- deterministic a priori error analysis

Thank you for interest!