Special Techniques for Kernel-Based Reconstruction of Functions from Meshless Data

Robert Schaback

schaback@math.uni-goettingen.de

Georg–August–Universität Göttingen

Overview

1. Scope of Symmetric Meshless Collocation
2. Complexity Reduction
3. Preconditioning

Part I

Scope of Symmetric Meshless Collocation

Hypotheses

- Linear mixed problem

$$L_i u = f_i \text{ in } \Omega_i \subset I\!R^d, \; i = 1, 2, \ldots,$$

- Existence and uniqueness of solution u in RKHS $\mathcal{U}$ with kernel Φ
- $L_i u$ and $L_i^x L_i^y \Phi(x, y)$ continuous on Ω_i

Result

Then the minimum norm solution by meshless symmetric collocation on dense subsets of Ω_i converges towards u.

Linear Problems

Given Problem

$$L_i u = f_i \text{ in } \Omega_i \subset I\!R^d,\ i = 1, 2, \ldots,$$

Discretize:

$$\begin{aligned} ((L_i u)(x_{ji}) &= f_i(x_{ji}) \quad x_{ji} \in \Omega_i \quad \text{for all } j \in I_i \\ (\delta_{x_{ji}} \circ L_i)\,(u) &= \delta_{x_{ji}} f_i \quad x_{ji} \in \Omega_i \quad \text{for all } j \in I_i \end{aligned}$$

Generally:

$$\lambda(u) = f_\lambda \in I\!R \text{ for all } \lambda \in \Lambda$$

Generalized Interpolation

Given $\Lambda \subset \mathcal{U}^*,\ f\ :\ \Lambda \to I\!R$

Find $u \in \mathcal{U}$ with

$$\lambda(u) = f(\lambda) \in I\!R \text{ for all } \lambda \in \Lambda$$

Minimum Norm Reconstruction

Given $\Lambda \subset \mathcal{U}^*, \ f \ : \ \Lambda \to I\!R$

Assume exact solution $u \in \mathcal{U}$ with

$$\lambda(u) = f(\lambda) \in I\!R \text{ for all } \lambda \in \Lambda$$

Reconstruction $\tilde{u} \in \mathcal{U}$ with

$$\lambda(\tilde{u}) = f(\lambda) \in I\!R \text{ for all } \lambda \in \Lambda$$

with minimal norm $\|\tilde{u}\|_{\mathcal{U}}$

Simplification: Hilbert Space

Closed subspaces

$$\mathcal{U}_\Lambda^\perp := \{u \in \mathcal{U} \ : \ \lambda(u) = 0 \text{ for all } \lambda \in \Lambda\}$$
$$\mathcal{U}_\Lambda := \left(\mathcal{U}_\Lambda^\perp\right)^\perp$$

Decompose solution $u = u_\Lambda + u_\Lambda^\perp \in \mathcal{U}_\Lambda + \mathcal{U}_\Lambda^\perp$

- Minimal norm solution is $\tilde{u} = u_\Lambda$
- Unique reconstruction iff $\mathcal{U}_\Lambda^\perp = \{0\}$

Problem: Construct u or $\tilde{u} = u_\Lambda$

Reproducing Kernel Hilbert Spaces

Space $\mathcal{U}$ of functions

Dual space $\mathcal{U}^*$ of functionals

Reproduction Property

$$\lambda(u) = (u, \lambda^x \Phi(x, \cdot))_{\mathcal{U}} \text{ for all } u \in \mathcal{U}, \quad \text{for all } \lambda \in \mathcal{U}^*$$

Minimum Norm Solution

Given $\Lambda_N = \{\lambda_1, \ldots, \lambda_N\} \subset \Lambda \subset \mathcal{U}^*,\ f\ :\ \Lambda \to I\!R$
Find $\tilde{u}_N \in \mathcal{U}$ of minimum norm $\|\tilde{u}_N\|_{\mathcal{U}}$ with

$$\lambda_j(\tilde{u}_N) = f(\lambda_j) \in I\!R \text{ for all } \lambda_j \in \Lambda_N$$

Solution:

- $\tilde{u}_N = \sum_k c_k \lambda_k^x \Phi(x, \cdot)$
- $\lambda_j(\tilde{u}_N) = \sum_k c_k \lambda_k^x \lambda_j^y \Phi(x, y) = f(\lambda_j)$

Used incrementally for Greedy Techniques

Problem: Usually $|\Lambda| = \infty$

Convergence

Given $\Lambda_N = \{\lambda_1, \ldots, \lambda_N\} \subset \mathcal{U}^*$ for $N \to \infty$

Sequence of minimum norm solutions $\tilde{u}_N \in \mathcal{U}$ is a Cauchy sequence, thus convergent.
Reason: $\|\tilde{u}_N\|^2_{\mathcal{U}} \nearrow \|\tilde{u}\|^2_{\mathcal{U}} \leq \|u\|^2_{\mathcal{U}}$ is a Cauchy sequence.
Construction works for $|\Lambda| = |I\!N|$.

Problem: Usually $|\Lambda| > |I\!N|$.

Density

Given $\Lambda_{\mathbb{N}} \subset \Lambda \subset \mathcal{U}^*, \ |\Lambda_{\mathbb{N}}| = |\mathbb{N}|$.
Construct minimal norm solution $\tilde{u}_{\mathbb{N}}$ for $\Lambda_{\mathbb{N}}$.
Problem: $\tilde{u}_{\mathbb{N}} = \tilde{u}$?

$$\lambda(\tilde{u}_{\mathbb{N}} - \tilde{u}) = 0 \text{ for all } \lambda \in \Lambda_{\mathbb{N}}$$

Assumption: $\Lambda_{\mathbb{N}}$ dense in Λ, i.e.

$$\begin{aligned} \lambda_j(v) &= 0 \quad \text{for all } \lambda_j \in \Lambda_{\mathbb{N}} \\ \Rightarrow \lambda(v) &= 0 \quad \text{for all } \lambda \in \Lambda \end{aligned}$$

Density and Continuity

Example:

$$\left(L_i(\tilde{u}_{\mathbb{N}} - \tilde{u})\right)(x_{ji}) = 0 \text{ for all } x_{ji} \in \Omega_i$$

Goal:

$$L_i(\tilde{u}_{\mathbb{N}}) = L_i(\tilde{u}) \text{ on } \Omega_i$$

if both continuous and $x_{ji} \in \Omega_i$ dense.

Final Result: Hypotheses

- Linear mixed problem

$$L_i u = f_i \text{ in } \Omega_i \subset I\!R^d,\ i = 1, 2, \ldots,$$

- Existence and uniqueness of solution u in RKHS $\mathcal{U}$ with kernel Φ
- $L_i u$ and $L_i^x L_i^y \Phi(x, y)$ continuous on Ω_i

Result

Then the minimum norm solution by meshless symmetric collocation on dense subsets of Ω_i converges towards u.

Part II

Complexity Reduction

Goal

Relaxing interpolation/collocation conditions yields a complexity reduction

Connection to

- machine learning
- support vector machines

Minimum Norm Reconstruction

Given $\Lambda \subset \mathcal{U}^*, \ f \ : \ \Lambda \to I\!R$
Reconstruction $\tilde{u} \in \mathcal{U}$ with

$$\lambda(\tilde{u}) = f(\lambda) \in I\!R \text{ for all } \lambda \in \Lambda$$

with minimal norm $\|\tilde{u}\|_{\mathcal{U}}$

Relaxed:

Given $\epsilon > 0$, reconstruct $\tilde{u}_\epsilon \in \mathcal{U}$ with

$$|\lambda(\tilde{u}_\epsilon) - f(\lambda)| \leq \epsilon \text{ for all } \lambda \in \Lambda$$

and minimal norm $\|\tilde{u}_\epsilon\|_{\mathcal{U}}$

Quadratic Optimization

Minimize $\|\tilde{u}_\epsilon\|^2_{\mathcal{U}}$ in RKHS $\mathcal{U}$ with kernel Φ
under linear constraints

$$-\epsilon \leq \lambda(\tilde{u}_\epsilon) - f(\lambda) \leq +\epsilon \text{ for all } \lambda \in \Lambda$$

Feasibility $\Rightarrow$ Solvability
Solution is **only** based on functions $\lambda^x \Phi(x, \cdot)$
where $\lambda \in \Lambda$ satisfies

$$\lambda(\tilde{u}_\epsilon) - f(\lambda) = \pm\epsilon$$

active constraint, **support vector**

Problem: How many support vectors are needed?

Learning

$$\begin{array}{ccc} \text{Stimulus space} & \overset{\text{Behavior}}{\longrightarrow} & \text{Response space} \\ S & \longrightarrow & R \\ x & \longmapsto & F(x) \end{array}$$

Training:
Given meshless pairs $(x_j, z_j) \in S \times R$
find F with $F(x_j) = z_j$ (Interpolation)
Then: use F on unknown inputs x

Learning with Kernels

“Kernel trick” for $R = I\!R$

$K : S \to \mathcal{F}$ = “feature” space

Find optimal $G : \mathcal{F} \to I\!R$ with

$$z_j = G(\underbrace{K(x_j)}_{y_j}) =: F(x_j)$$

Interpolation in RKHS on $\mathcal{F}$ with kernel Φ

Solution:

$$\begin{aligned} G(y) &= \sum_j \alpha_j \Phi(y_j, y) \\ F(x) = G(K(x)) &= \sum_j \alpha_j \Phi(K(x_j), K(x)) \end{aligned}$$

Optimized Learning with Kernels

Minimize $\|G\|_{\mathcal{U}}^2$ under constraints

$$-\epsilon \leq z_j - G(K(x_j)) \leq \epsilon$$

Feasibility $\Rightarrow$ Solvability

Solution is **only** based on training samples x_j

where G satisfies

$$G(K(x_j)) - z_j = \pm\epsilon$$

active constraint, **support vector**

Example: Michael Schumacher

Example: Learning the peaks function of MATLAB

Example: Census data

Connection to Greedy Methods

$\Lambda_N = \{\lambda_1, \ldots, \lambda_N\} \subset \Lambda \subset \mathcal{U}^*, \ f \ : \ \Lambda \rightarrow I\!R, \ \epsilon > 0$.

Assume $\tilde{u}_N \in \mathcal{U}$ is minimum norm function with

$$\lambda_j(\tilde{u}_N) = f(\lambda_j) \text{ for all } \lambda_j \in \Lambda_N.$$

Then find $\lambda \in \Lambda \setminus \Lambda_N$ such that

$$\epsilon_N := \max_{\lambda} |\lambda(\tilde{u}_N) - f(\lambda)|$$

If $\epsilon_N > \epsilon$ set $\lambda_{N+1} := \lambda$ and repeat.

Under weak assumptions: $\lim_{n\to\infty} \epsilon_N = 0$

Example: Greedy reconstruction of the peaks function

Open Problem

Within quadratic optimization:

Given $\Lambda \subset \mathcal{U}^*,\ f\ :\ \Lambda \to I\!R$, given $\epsilon > 0$

Determine $N(\epsilon)$=number of support vectors

Part III

Preconditioning

Polynomials on Meshless Data

$$
\begin{aligned}
\mathbb{P}_m^d &:= d\text{–variate poly's of order } \leq m \\
&= \text{span}\,\{p_1,\ldots,p_Q\},\ Q = \binom{m-1+d}{d} \\
\Lambda &:= \{\lambda_1,\ldots,\lambda_N\} \subset \mathcal{U}^* \\
P_m(\Lambda) &:= (\lambda_k(p_j))_{1\leq j\leq Q,1\leq k\leq N} \\
\mu(\Lambda) &:= \max\{m\ :\ \text{rank}\, P_m(\Lambda) < N\} \\
\text{rank}\, P_{\mu(\Lambda)}(\Lambda) &< N = \text{rank}\, P_{\mu(\Lambda)+1}(\Lambda)
\end{aligned}
$$

Divided Differences

$$\begin{aligned} D_m(\Lambda) &:= \text{ker } P_m(\Lambda) \subseteq I\!R^N \\ \alpha \in D_m &\Leftrightarrow \sum_{k=1}^{N} \alpha_k \lambda_k(p) = 0 \text{ for all } p \in P_m^d \\ I\!R^N = D_0(\Lambda) &\supseteq D_1(\Lambda) \supseteq \ldots \\ \ldots &\supseteq D_{\mu(\lambda)}(\Lambda) \neq \{0\} = D_{\mu(\lambda)+1}(\Lambda) \end{aligned}$$

Divided difference basis

$\alpha^1, \ldots, \alpha^N \in I\!R^N$ for X

$$\alpha^j \in D_{t_j} \setminus D_{t_j+1}, \; 0 = t_1 \leq \ldots \leq t_N = \mu(\Lambda)$$

Divided Difference Matrix

$$A = \left(\alpha_k^j\right)_{1 \le j,k \le N}$$

with divided difference basis $\alpha^1, \ldots, \alpha^N \in I\!R^N$ wrt. Λ

$$\sum_{k=1}^{N} \alpha_k^j \lambda_k(p) = 0 \text{ for all } p \in I\!P_{t_j}^d,\ 1 \le j \le N$$

for polynomial orders $0 = t_1 \le \ldots \le t_N = \mu(\Lambda)$

Depends on Λ, **not on any rbf**

Scaled analytic radial basis function

$$\phi_c(r) = f(c^2r^2) = \sum_{\ell=0}^{\infty} f_\ell c^{2\ell} r^{2\ell}$$

Interpolation matrix:

$$C_c := \left(\lambda_k^x \lambda_j^y \phi(c^2\|x-y\|_2)\right)_{1\le j,k\le N}$$

Scaling matrix:

$$B_c := \left(c^{-t_j}\delta_{jk}\right)_{1\le j,k\le N}$$

Preconditioning Transformation

$$e_j^T \cdot B_c^T \cdot A \cdot C_c \cdot A \cdot B_c \cdot e_k$$

$$= \sum_{r=1}^{j} c^{-t_j} \alpha_r^j \sum_{s=1}^{k} c^{-t_k} \alpha_s^k \sum_{\ell=0}^{\infty} f_\ell c^{2\ell} \lambda_r^x \lambda_s^y \|x-y\|_2^{2\ell}$$

$$= \sum_{2\ell \geq t_j+t_k}^{\infty} f_\ell c^{2\ell-t_j-t_k} \underbrace{\sum_{r=1}^{j} \sum_{s=1}^{k} \alpha_r^j \alpha_s^k \lambda_r^x \lambda_s^y \|x-y\|_2^{2\ell}}_{\text{degree } \leq 2\ell-t_j-t_k}$$

$$\to f_{(t_j+t_k)/2} \sum_{r=1}^{j} \sum_{s=1}^{k} \alpha_r^j \alpha_s^k \lambda_r^x \lambda_s^y \|x-y\|_2^{t_j+t_k}$$

for $t_j + t_k = 2\ell$ even

Limit Matrix

$$\lim_{c\to 0} B_c^T \cdot A \cdot C_c \cdot A \cdot B_c$$

converges to a matrix with entries

$$\begin{cases} 0 & t_j + t_k \text{ odd} \\ f_{(t_j+t_k)/2} \sum_{r=1}^{j} \sum_{s=1}^{k} \alpha_r^j \alpha_s^k \lambda_r^x \lambda_s^y \|x-y\|_2^{t_j+t_k} & t_j + t_k = 2\ell \text{ even} \end{cases}$$

Example: Preconditioning

Least Interpolation

Using the polynomial basis

$$v_j(x) := \sum_{k=1}^{N} \alpha_k^j \lambda_k^y \|x - y\|_2^{2t_j},\ 1 \le j \le N$$

one can solve the generalized polynomial interpolation problem

$$\lambda_j(p) = \lambda_j(u),\ 1 \le j \le N$$

and this interpolation uses the least possible degree $\mu(\Lambda)$.

Open Problems

- Use for real–world problems
- Localize

The End