Learning methods for solving PDEs

**ANNs**


Model reduction and neural networks for parametric PDEs.

**GPs**

Multigrid with rough coefficients and Multiresolution operator decomposition from Hierarchical Information Games. H. Owhadi, SIREV, 2017

**Time dependent**: Numerical Gaussian processes for time-dependent and nonlinear partial differential equations M Raissi, P Perdikaris, GE Karniadakis, SISC 2018

**RBF collocation methods**: R. Schaback and H. Wendland, 2006
**Interplays with numerical approximation**: Sard, Larkin, Diaconis, Suldin, Kimeldorf and Wahba

GPs: *More theoretically well-founded and with a long history of interplays with numerical approximation but were limited to linear/quasi-linear/time-dependent PDEs*
Generalization of GP methods to arbitrary nonlinear PDEs


Properties

- Provably convergent for forward problems
- Interpretable and amenable to numerical analysis
- Solve forward and inverse problems
- Inherit the complexity of SOA solvers for dense kernel matrices
- Could be used to develop a theoretical understanding of PINNs
A simple prototypical non-linear PDE

\[
-\Delta u^\dagger + \tau(u^\dagger) = f, \quad x \in \Omega,
\]

\[
u^\dagger = g, \quad x \in \partial\Omega,
\]

\(f: \Omega \to \mathbb{R}, \ g: \partial\Omega \to \mathbb{R}\) and \(\tau: \mathbb{R} \to \mathbb{R}\): given continuous functions.

\(\tau\): Such that the PDE has a unique strong solution

Generalizes to arbitrary non-linear PDEs
\[
\begin{cases}
-\Delta u^\dagger + \tau(u^\dagger) = f, & x \in \Omega, \\
u^\dagger = g, & x \in \partial\Omega,
\end{cases}
\]

The method

\[K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}: \text{Given kernel.}\]

\[X_1, \ldots, X_N: \text{Collocation points on } \Omega \text{ and } \partial\Omega\]

Approximate \( u^\dagger \) with the minimizer \( u \) of

\[
\begin{align*}
\text{Minimize} & \quad \|u\|^2_K \\
\text{subject to} & \quad -\Delta u(X_i) + \tau(u(X_i)) = f(X_i), \quad X_i \in \Omega, \\
\text{and} & \quad u(X_i) = g(X_i), \quad X_i \in \partial\Omega,
\end{align*}
\]
Theorem

Assume that

- $K$ is chosen so that
  - $\mathcal{H} \subset H^s(\Omega)$ for some $s > s^*$, where $s^* = \frac{d}{2} + \text{order of PDE}$ (order of PDE = 2)
  - $u^\dagger \in \mathcal{H}$
  - Fill distance of $\{X_1, \ldots, X_N\}$ goes to zero as $N \to \infty$

Then, as $N \to \infty$

- $u \to u^\dagger$ pointwise in $\bar{\Omega}$
- $u \to u^\dagger$ in $H^t(\Omega)$ for $t < s$

$\mathcal{H}$: RKH space defined by kernel $K$
Implementation

Minimize $\|u\|_K^2$
subject to $-\Delta u(X_i) + \tau(u(X_i)) = f(X_i), \ X_i \in \Omega,$
and $u(X_i) = g(X_i), \ X_i \in \partial \Omega,$

\[\begin{aligned}
\min_{z^{(1)}, z^{(2)}} & \quad \min_u \|u\|_K^2 \\
\text{s.t.} & \quad u(X_i) = z_i^{(1)} \quad \text{and} \quad -\Delta u(X_i) = z_i^{(2)} \\
& \quad z_i^{(2)} + \tau(z_i^{(1)}) = f(X_i) \text{ for } X_i \in \Omega \\
& \quad z_i^{(1)} = g(X_i) \text{ for } X_i \in \partial \Omega
\end{aligned}\]

Reduction theorem

$z = (z^{(1)}, z^{(2)})$
$\phi = (\phi^{(1)}, \phi^{(2)})$
$
\phi_i^{(1)} = \delta_{X_i}$
$
\phi_i^{(2)} = \delta_{X_i} \circ \Delta$

$u(x) = K(x, \phi)K(\phi, \phi)^{-1}z$

\[\begin{aligned}
\min_{z^{(1)}, z^{(2)}} & \quad z^T K(\phi, \phi)^{-1}z \\
& \quad z_i^{(2)} + \tau(z_i^{(1)}) = f(X_i) \text{ for } X_i \in \Omega \\
& \quad z_i^{(1)} = g(X_i) \text{ for } X_i \in \partial \Omega
\end{aligned}\]

$(K(x, \phi))_i = \int K(x, y)\phi_i(y) \, dy$

$(K(\phi, \phi))_{i,j} = \int \phi_i(x)K(x, y)\phi_j(y) \, dx \, dy$
\[
\begin{aligned}
\min_{z(1), z(2)} & \quad z^T K(\phi, \phi)^{-1} z \\
& \quad z_i^{(2)} + \tau(z_i^{(1)}) = f(X_i) \text{ for } X_i \in \Omega \\
& \quad z_i^{(1)} = g(X_i) \text{ for } X_i \in \partial \Omega \\
\end{aligned}
\]

Eliminate \(z^{(2)}\)

\[
\begin{aligned}
\min_{z^{(1)}} & \quad (z_i^{(1)}, g(X_i), f(X_i) - \tau(z_i^{(1)}))^T K(\phi, \phi)^{-1} (z_i^{(1)}, g(X_i), f(X_i) - \tau(z_i^{(1)})) \\
\end{aligned}
\]

Gauss-Newton Iteration

\[
\begin{aligned}
 z_i^{(1), n+1} &= z_i^{(1), n} + \delta z_i^{(1), n} \\
\min_{\delta z^{(1)}} & \quad Z^T K(\phi, \phi)^{-1} Z \\
\end{aligned}
\]

\[
\begin{aligned}
 Z &= (z_i^{(1), n} + \delta z_i^{(1), n}, g(X_i), f(X_i) - \tau(z_i^{(1), n}) - \delta z_i^{(1), n} \nabla \tau(z_i^{(1), n})) \\
\end{aligned}
\]

Converges in 2 to 7 steps

Inherits the complexity of fast linear solvers for \(K(\phi, \phi)\)

[Schäfer, Katzfuss and O., 2020]: \(O(N \log^{2d}(\frac{N}{\varepsilon}))\) complexity
Gauss-Newton Iteration $\longleftrightarrow$ Successive linearization of the PDE

\[
\begin{cases}
-\Delta u^\dagger + \tau(u^\dagger) = f, & x \in \Omega, \\
u^\dagger = g, & x \in \partial\Omega,
\end{cases}
\]

\[u^{n+1} = u^n + \delta u^n\]

Given $u^n$ solve for $\delta u^n$

\[
\begin{cases}
-\Delta(u^n + \delta u^n) + \tau(u^n) + \delta u^n \nabla \tau(u^n) = f, & x \in \Omega, \\
u^n + \delta u^n = g, & x \in \partial\Omega,
\end{cases}
\]
Numerical experiments

\[ K(x, x') = \exp \left( - \frac{|x - x'|^2}{\sigma^2} \right) \]

FD: Finite difference
\[ \partial_t u + u \partial_s u - \nu \partial_s^2 u = 0, \quad \forall (s, t) \in [-1, 1] \times [0, \infty), \]
\[ u(s, 0) = -\sin(\pi x), \]
\[ u(-1, t) = u(1, t) = 0. \]

\[ K((x, t), (x', t')) = \exp(-20|x - x'|^2 - 3|t - t'|^2) \]

<table>
<thead>
<tr>
<th>( N )</th>
<th>600</th>
<th>1200</th>
<th>2400</th>
<th>4800</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L^2 ) error</td>
<td>1.75e-02</td>
<td>7.90e-03</td>
<td>8.65e-04</td>
<td>9.76e-05</td>
</tr>
<tr>
<td>( L^\infty ) error</td>
<td>6.61e-01</td>
<td>6.39e-02</td>
<td>5.50e-03</td>
<td>7.36e-04</td>
</tr>
</tbody>
</table>
\[
\begin{align*}
\| \nabla u(x) \|^2 &= f(x)^2 + \epsilon \Delta u(x), \quad \forall x \in \Omega, \\
u(x) &= 0, \quad \forall x \in \partial \Omega,
\end{align*}
\]

\[K(x, x') = \exp\left(-\frac{|x - x'|^2}{\sigma^2}\right)\]

<table>
<thead>
<tr>
<th></th>
<th>1200</th>
<th>1800</th>
<th>2400</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L^2$ error</td>
<td>3.7942e-04</td>
<td>1.3721e-04</td>
<td>1.2606e-04</td>
<td>1.1025e-04</td>
</tr>
<tr>
<td>$L^\infty$ error</td>
<td>5.5768e-03</td>
<td>1.4820e-03</td>
<td>1.3982e-03</td>
<td>9.5978e-04</td>
</tr>
</tbody>
</table>

Collocation points

Loss function history

Contour of errors
Inverse Problem

\[
\begin{cases}
- \text{div} (\exp(a) \nabla u)(x) = f(x), & x \in \Omega, \\
 u(x) = 0, & x \in \partial \Omega.
\end{cases}
\]

\(a, u\): Unknown. \(u\) observed at pink points.
Problem: Recover \(a\) and \(u\).

Minimize \[\|u\|^2_K + \|a\|^2_1\]
subject to \[- \text{div} (\exp(a) \nabla u)(X_i) = f(X_i), & X_i \in \Omega, \]
and \[u(X_i) = Y_i, \quad (X_i, Y_i) \text{ is data point}, \]
and \[u(X_i) = 0, \quad X_i \in \partial \Omega, \]
Inverse Problem

\[
\begin{cases}
- \text{div} \left( \exp(a) \nabla u \right)(x) = f(x), & x \in \Omega, \\
u(x) = 0, & x \in \partial\Omega.
\end{cases}
\]

\(a, u\): Unknown. \(u\) observed at pink points.

Problem: Recover \(a\) and \(u\).