高斯过程求解偏微分方程

黄政宇
北京大学北京国际数学研究中心
北京大学国际机器学习研究中心

本堂课大纲

- 求解偏微分方程问题
  - 核岭插值
  - 随机特征方法

- 含参数偏微分方程问题
  - 随机特征方法
  - 高斯回归

求解偏微分方程问题

- 偏微分方程问题

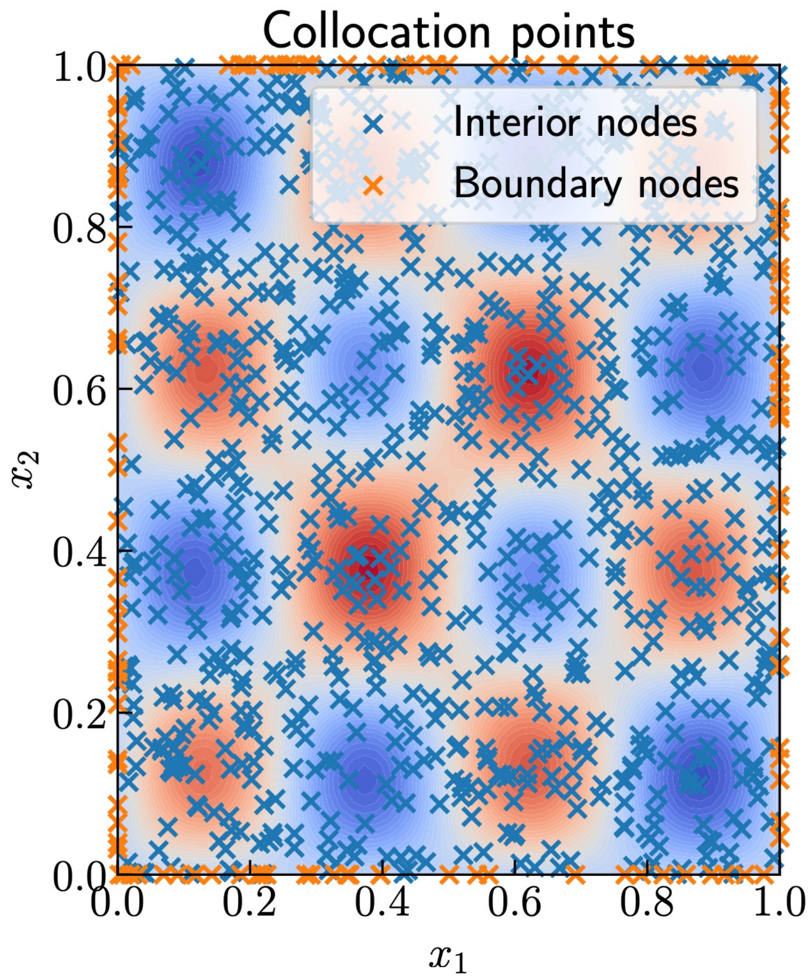
$$\mathcal{L}(x,u)=0 \qquad x\in\Omega$$
$$\mathcal{B}(x,u)=0 \qquad x\in\partial\Omega$$

- 核岭插值

$$\hat{u}=\mathrm{argmin}_{u\in\mathcal{H}_\kappa}\|u\|_{\mathcal{H}_\kappa}$$
$$\mathcal{L}(x_i,u(x_i))=0 \qquad x_i\in X^{int}$$
$$\mathcal{B}(x_i,u(x_i))=0 \qquad x_i\in X^{bd}$$

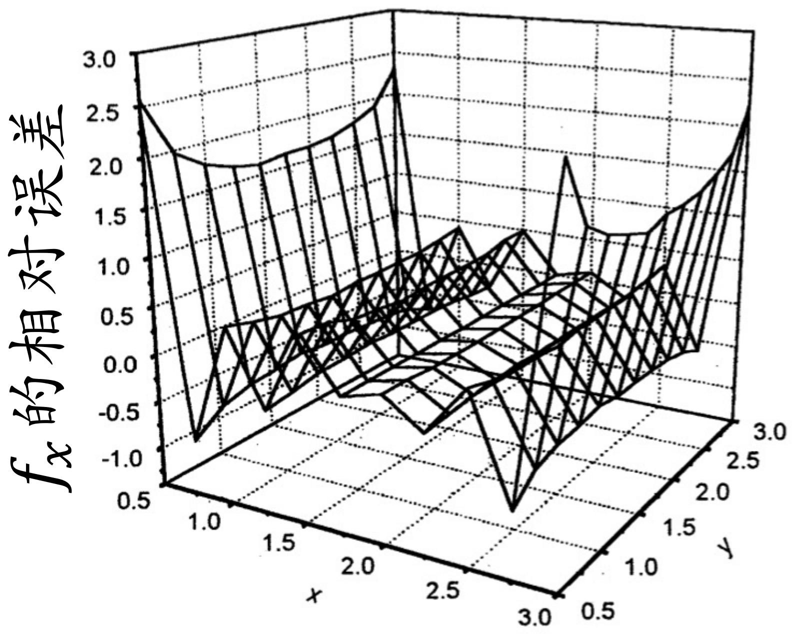
简单的想法：

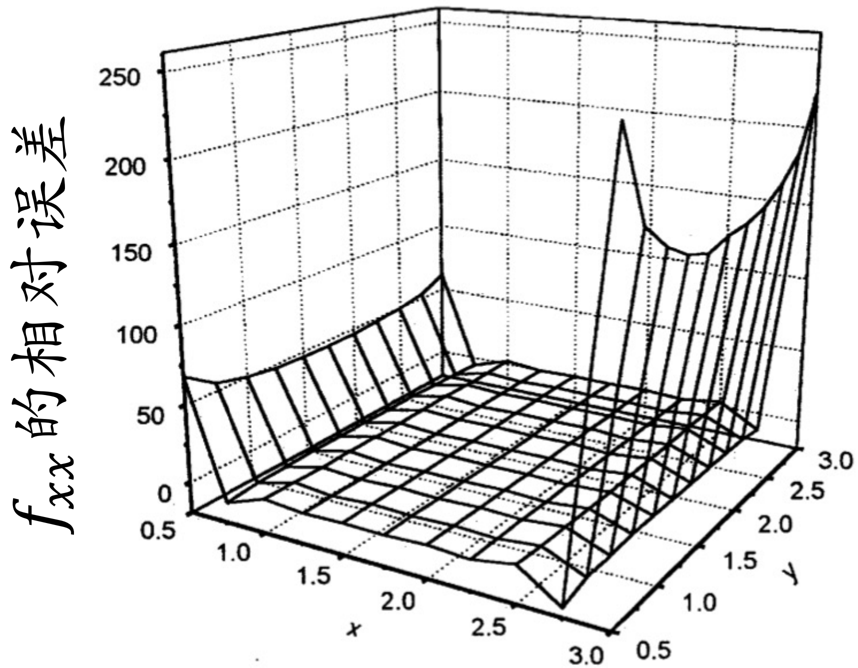
基底函数 $\kappa(x,x_i)\ \ i=1,2,\cdots,n$



插值问题

- 练习
  - 无噪音2维训练数据$\{(x_i, f(x_i)) | i = 1,2\cdots\cdots n\}$
  - $f(x_i) = \sin(x)\cos(y) \qquad x \in [0.5,3.0], y \in [0.5,3.0]$
  - x_i为10×10均匀格点





插值问题

插值问题

- 无噪音训练数据 $\{(x_i, f(x_i), \nabla_x f(x_i), \Delta f(x_i)) | i = 1,2 \cdots \cdots n\}$

核岭插值

$$\hat{f} = \operatorname{argmin}_{f \in \mathcal{H}_k} \|f\|_{\mathcal{H}_k}$$

$$\text{使得} \quad f(x_i) = y_i \qquad \nabla_x f(x_i) = y_i' \qquad \Delta f(x_i) = y_i''$$

$$\hat{f}(x) = \sum_{i=1}^{n} \alpha_i \kappa(x, x_i) + \sum_{i=1}^{n} \alpha_i' \nabla_x \kappa(x, x_i) + \sum_{i=1}^{n} \alpha_i'' \Delta_x \kappa(x, x_i)$$

基底函数 $\kappa(x, x_i), \nabla_x \kappa(x, x_i), \Delta_{xx} \kappa(x, x_i), i = 1,2, \cdots, n$

插值问题

- 对于足够光滑的平稳核函数

$$\frac{\partial^k \kappa(x,y)}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_k}} \in \mathcal{H}_\kappa$$

且

$$\left\langle \frac{\partial^k \kappa(x,y)}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_k}}, f(y) \right\rangle = (-1)^k \frac{\partial^k f(x)}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_k}}$$

基本想法：有限差分逼近

本章后续，我们均考虑足够光滑的平稳核函数

核岭插值

- 线性偏微分方程问题

$$
\begin{aligned}
-\Delta u &= f(x) \qquad x \in \Omega \\
u &= \bar{u} \qquad\;\; x \in \partial\Omega
\end{aligned}
$$

- 核岭插值

$$\hat{u} = \operatorname{argmin}_{u \in \mathcal{H}_k} \|u\|_{\mathcal{H}_k}$$

偏微分方程约束 $\quad -\Delta u(x_i) = f(x_i) \qquad x_i \in X^{int}$

边界条件约束 $\quad u(x_i) = \bar{u}(x_i) \qquad x_i \in X^{bd}$

核岭插值

- 核岭插值

$$\hat{u} = \operatorname{argmin}_{u\in\mathcal{H}_k}\|u\|_{\mathcal{H}_k}$$

使得 $u(x_i) = \bar{u}(x_i) \quad x_i \in X^{bd}, \quad -\Delta u(x_i) = f(x_i) \quad x_i \in X^{int}$

- 表示定理

$$u(x) = \sum_{x_i\in X^{bd}} \alpha_i \kappa(x, x_i) + \sum_{x_i\in X^{int}} \alpha_i'' \Delta_x \kappa(x, x_i)$$

- 求解

$$k_{xX} = \left[\kappa\left(x, X^{bd}\right) \ \Delta_y \kappa\left(x, X^{int}\right)\right]$$

$$K_{XX} = \begin{bmatrix} \kappa\left(X^{bd}, X^{bd}\right) & \Delta_y \kappa\left(X^{bd}, X^{int}\right) \\ \Delta_x \kappa\left(X^{int}, X^{bd}\right) & \Delta_x \Delta_y \kappa\left(X^{int}, X^{int}\right) \end{bmatrix}$$

$$f = [\cdots \bar{u}(x_i) \cdots - f(x_i) \cdots]^T$$

$$u(x) = k_{xX} K_{XX}^{-1} f$$

核岭插值

- 非线性偏微分方程问题

$$\begin{aligned} -\Delta u + \tau(u) = f(x) \quad & x \in \Omega \\ u = \bar{u} \quad & x \in \partial\Omega \end{aligned}$$

- 两层优化

$$\hat{u} = \operatorname{argmin}_{u \in \mathcal{H}_k} \|u\|_{\mathcal{H}_k}$$

$$\begin{cases} \underset{z=(z^{bd}, z^{int}, z_\Delta^{int})}{\text{minimize}} \begin{cases} \operatorname{argmin}_{u \in \mathcal{H}_k} \|u\|_{\mathcal{H}_k} \\ u\left(X^{bd}\right) = z^{bd}, \\ u\left(X^{int}\right) = z^{int}, \\ \Delta u\left(X^{int}\right) = z_\Delta^{int} \end{cases} \\ -z_\Delta^{int} + \tau\left(z^{int}\right) = f(X^{int}) \\ z^{bd} = \bar{u}(X^{bd}) \end{cases}$$

核岭插值

- 内层优化

$$\hat{u} = \mathrm{argmin}_{u\in\mathcal{H}_k}\|u\|_{\mathcal{H}_k}$$

$$使得 u\left(X^{bd}\right) = z^{bd}, u\left(X^{int}\right) = z^{int}, \Delta u\left(X^{int}\right) = z_{\Delta}^{int}$$

- 表示定理

$$u(x) = \sum_{x_i\in X^{bd},X^{int}} \alpha_i \kappa(x, x_i) \ + \sum_{x_i\in X^{int}} \alpha_i'' \Delta_x \kappa(x, x_i)$$

- 求解

$$k_{xX} = \left[\kappa\left(x, X^{bd}\right) \ \kappa\left(x, X^{int}\right) \ \Delta_y\kappa\left(x, X^{int}\right)\right]$$

$$K_{XX} = \begin{bmatrix} \kappa\left(X^{bd}, X^{bd}\right) & \kappa\left(X^{bd}, X^{int}\right) & \Delta_y\kappa\left(X^{bd}, X^{int}\right) \\ \kappa\left(X^{int}, X^{bd}\right) & \kappa\left(X^{int}, X^{int}\right) & \Delta_y\kappa\left(X^{int}, X^{int}\right) \\ \Delta_x\kappa\left(X^{int}, X^{bd}\right) & \Delta_y\kappa\left(X^{int}, X^{int}\right) & \Delta_x\Delta_y\kappa\left(X^{int}, X^{int}\right) \end{bmatrix}$$

$$u(x) = k_{xX}K_{XX}^{-1}z \qquad \|u\|_{\mathcal{H}_k} = z^T K_{XX}^{-1} z$$

核岭插值

- 外层优化

$$\underset{z=(z^{bd},z^{int},z_{\Delta}^{int})}{\text{minimize}} \quad z^T K_{XX}^{-1} z$$

$$-z_{\Delta}^{int} + \tau\left(z^{int}\right) = f\left(X^{int}\right), \quad z^{bd} = \bar{u}(X^{bd})$$

- 两层优化

$$\underset{z=(z^{bd},z^{int},z_{\Delta}^{int})}{\text{minimize}} \quad z^T K_{XX}^{-1} z$$

$$F(z) = y$$

Newton法：$z^k \rightarrow z^{k+1}$

$$\underset{z=(z^{bd},z^{int},z_{\Delta}^{int})}{\text{minimize}} \quad z^T K_{XX}^{-1} z$$

$$F\left(z^k\right) + F\left(z^k\right)(z - z^k) = y$$

Bochner定理

一个定义在R^N中的有平移不变性的实值核函数，那么当κ能表示为

$$\kappa(\tau) = \int_{R^N} e^{2\pi i\omega\cdot\tau} p(\omega) d\omega$$

其中$\mu(\omega)$是有限正测度，我们考虑特殊情况$d\mu(\omega) = p(\omega)d\omega$。

定义$\psi(x,\omega) = e^{2\pi i\omega\cdot x}$，那么

$$\kappa(x - x') = \int_{R^N} e^{2\pi i\omega\cdot(x-x')} p(\omega) d\omega$$
$$= \mathbb{E}_{\omega\sim p(\omega)}[\psi(x,\omega)\psi^*(x',\omega)]$$

随机特征方法

- 偏微分方程问题

$$\mathcal{L}(x,u)=0 \qquad x\in\Omega$$
$$\mathcal{B}(x,u)=0 \qquad x\in\partial\Omega$$

- 随机特征方法

$$\hat{u}(x)=\sum_{i=1}^{D}c_i\psi(x,\omega_i)$$

$$\min \sum_{x_i\in X^{int}}\lambda_{Ii}\left\|\mathcal{L}(x_i,u(x_i))\right\|^2+\sum_{x_i\in X^{bd}}\lambda_{Bi}\left\|\mathcal{B}(x_i,u(x_i))\right\|^2+\lambda\|c\|_2^2$$

二次优化问题

求解偏微分方程问题

- 传统方法(有限体积、有限元方法)
  - 需要有网格
  - 局部基底函数
  - 能保持守恒律
- 基于高斯回归的方法（核岭回归、随机特征方法）
  - 不需要有网格、只需要采样点
  - 全局基底函数
  - 对足够光滑的问题，能达到指数收敛（谱方法）
  - 超参选取、可能有数值不稳定

本堂课大纲

- 求解偏微分方程问题
  - 核岭插值
  - 随机特征方法

- 含参数偏微分方程问题
  - 随机特征方法
  - 高斯回归

含参数偏微分方程问题

- 偏微分方程问题

$$\mathcal{L}(x,u,a)=0 \quad x\in\Omega$$
$$\mathcal{B}(x,u)=0 \quad x\in\partial\Omega$$

$$\mathcal{G}^{\dagger}: a\to u$$

- 代理模型(surrogate model)

两个函数空间 $\mathcal{A}=\{a:\Omega\mapsto\mathbb{R}\}$ $\mathcal{U}=\{u:\Omega\mapsto\mathbb{R}\}$，求解偏微分方程给出映射 $\mathcal{G}^{\dagger}:\mathcal{A}\mapsto\mathcal{U}$

降阶模型$\mathcal{G}_{\theta}\approx\mathcal{G}^{\dagger}$是一个这样的映射的近似

算子学习(OPERATOR LEARNING)

- 偏微分方程问题

$$\mathcal{L}(x,u,a)=0 \quad x\in\Omega$$
$$\mathcal{B}(x,u)=0 \quad x\in\partial\Omega$$

$$\mathcal{G}^{\dagger}: a\to u$$

- 代理模型(surrogate model)

两个函数空间 $\mathcal{A}=\{a:\Omega\mapsto\mathbb{R}\}$ $\mathcal{U}=\{u:\Omega\mapsto\mathbb{R}\}$，求解偏微分方程给出映射 $\mathcal{G}^{\dagger}:\mathcal{A}\mapsto\mathcal{U}$

给定数据 $\{a_i,u_i=\mathcal{G}^{\dagger}(a_i)\}$

寻找近似$\mathcal{G}_{\theta}\approx\mathcal{G}^{\dagger}$，比如降阶模型。

再生核希尔伯特空间

- 核函数

  函数$f: R^N \to R$

  核函数$\kappa: R^N \times R^N \to R$

  再生核希尔伯特空间$\mathcal{H}_\kappa$:

$$\mathcal{H}_\kappa = \overline{\mathrm{span}\{\kappa(\cdot, x_i) c_i\}}$$

  表示定理:

$$f = \sum_{i=1} \kappa(\cdot, x_i) c_i \quad (c_i \in R)$$

  $\kappa(\cdot, x): \mathcal{H}_\kappa \to R$定义了有界线性映射

$$f \to \langle \kappa(\cdot, x), f \rangle_{\mathcal{H}_\kappa} = f(x)$$

再生核希尔伯特空间

➤ 算子值(Operator valued)核函数

算子$\mathcal{G}^{\dagger}:\mathcal{A}\rightarrow\mathcal{U}$，$\mathcal{A}$和$\mathcal{U}$是两个实希尔伯特空间

核函数$\kappa:\mathcal{A}\times\mathcal{A}\mapsto\mathcal{L}(\mathcal{U},\mathcal{U})$其中$\mathcal{L}(\mathcal{U},\mathcal{U})$表示在空间$\mathcal{U}$上的所有有界线性算子构成的 Banach 空间

正定性：

$$\sum_{i,j=1}^{n}\langle u_i,\kappa(a_i,a_j)u_j\rangle_{\mathcal{U}}\geq 0$$

再生性：

$\langle u,f(a)\rangle_{\mathcal{U}}=\langle\kappa(\cdot,a)u,f(\cdot)\rangle_{\mathcal{H}_\kappa}\forall a\in\mathcal{A},u\in\mathcal{U},f:\mathcal{A}\rightarrow\mathcal{U}$

表示定理：$f=\sum_{i=1}\kappa(\cdot,a_i)u_i$

再生核希尔伯特空间：

$$\mathcal{H}_\kappa=\overline{\mathrm{span}\{\kappa(\cdot,a_i)u_i\}}$$

再生核希尔伯特空间

➢ 代理模型: 算子学习

给定数据$\{a_i, u_i = \mathcal{G}^\dagger(a_i)\}$，以及再生核希尔伯特空间$\mathcal{H}_\kappa$，求解核岭回归（插值）

$$\hat{\mathcal{G}} = \mathrm{argmin}_{\mathcal{G}\in\mathcal{H}_\kappa}\|\mathcal{G}\|_{\mathcal{H}_\kappa} \text{ 使得 } \mathcal{G}(a_i) = u_i$$

➢ 表示定理：

$$\hat{\mathcal{G}}(\cdot) = \sum_{i=1}\kappa(\cdot, a_i)y_i : \quad \mathcal{A} \rightarrow \mathcal{U}$$

$\kappa(\cdot, a_i)$:难以构造

- 算子值(Operator valued)核函数

假设

$$\kappa(a,a') = \int \varphi(a,\omega) \otimes \varphi(a',\omega)\rho(\omega)d\omega$$

其中

$$\varphi(a,\omega): \mathcal{A}\times\Theta \to \mathcal{U}$$
$$(u_1 \otimes u_2)u_3 = \langle u_2, u_3\rangle_{\mathcal{U}} u_1$$

我们有

$$\kappa(a,a')u = \int \langle\varphi(a',\omega), u\rangle_{\mathcal{U}} \varphi(a,\omega)\,\rho(\omega)d\omega \in \mathcal{U}$$

随机特征方法

➢ 算子值(Operator valued)核函数

再生性：

$$
\begin{aligned}
\langle \mathcal{G}(a), u\rangle_{\mathcal{U}} &= \langle \mathcal{G}(\cdot),\ \kappa(\cdot, a)u\rangle_{\mathcal{H}_\kappa} \\
&= \left\langle \mathcal{G}(\cdot),\ \int \varphi(\cdot,\omega) \otimes \varphi(a,\omega)\rho(\omega)d\omega\, u \right\rangle_{\mathcal{H}_\kappa} \\
&= \int \langle \varphi(a,\omega), u\rangle_{\mathcal{U}} \langle \varphi(\cdot,\omega), \mathcal{G}(\cdot)\rangle_{\mathcal{H}_\kappa} \rho(\omega)d\omega \\
&= \left\langle \int c_f(\omega)\varphi(a,\omega)\rho(\omega)d\omega\, , u \right\rangle_{\mathcal{U}}
\end{aligned}
$$

其中 $c(\omega) = \langle \varphi(\cdot,\omega), \mathcal{G}(\cdot)\rangle_{\mathcal{H}_\kappa}$，我们有

$$
\mathcal{G}(a) = \int c(\omega)\varphi(a,\omega)\rho(\omega)d\omega \approx \frac{1}{D}\sum_{i=1}^{D} c_i \varphi(a,\omega_i)
$$

随机特征方法

- 核岭回归

$$\hat{\mathcal{G}} = \operatorname*{argmin}_{\mathcal{G}\in\mathcal{H}_\kappa} \frac{1}{n}\sum_{i=1}^{n}(\mathcal{G}(a_i) - u_i)^2 + \lambda\|\mathcal{G}\|_{\mathcal{H}_\kappa}^2$$

- 随机特征方法

$$\mathcal{G}(a) = \frac{1}{D}\sum_{i=1}^{D} c_i\varphi(a, \omega_i)$$

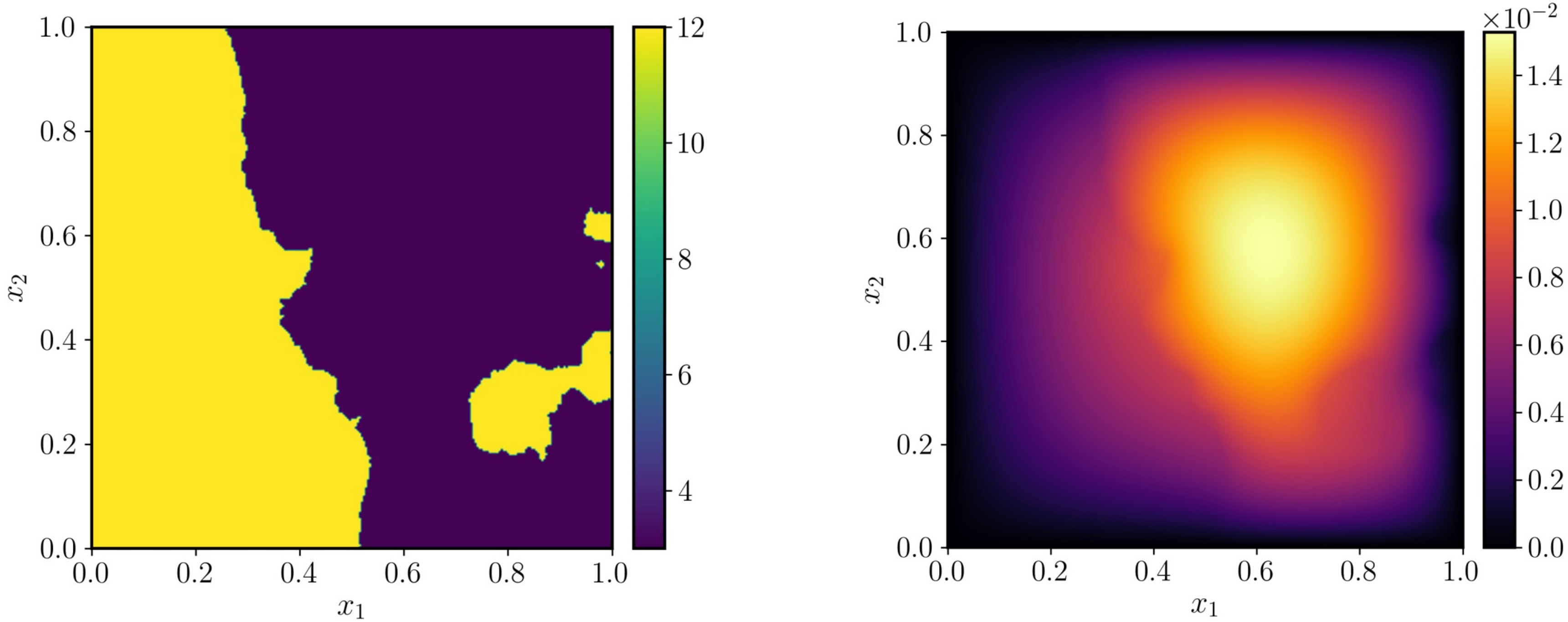
$$c = \operatorname*{argmin}_{c\in R^D} \sum_{i=1}^{n}\frac{1}{n}\left\|\frac{1}{D}\sum_{l=1}^{D} c_l\varphi(a_i, \omega_l) - u_i\right\|_{\mathcal{U}}^2 + \frac{\lambda}{D}\|c\|_2^2$$

例子

- Darcy 流方程

$$
\begin{aligned}
-\nabla \cdot (a(x)\nabla u) &= f(x), \qquad && x \in D \\
u(x) &= 0 && x \in \partial D
\end{aligned}
$$

$$\mathcal{G}^{\dagger} : a \mapsto u$$



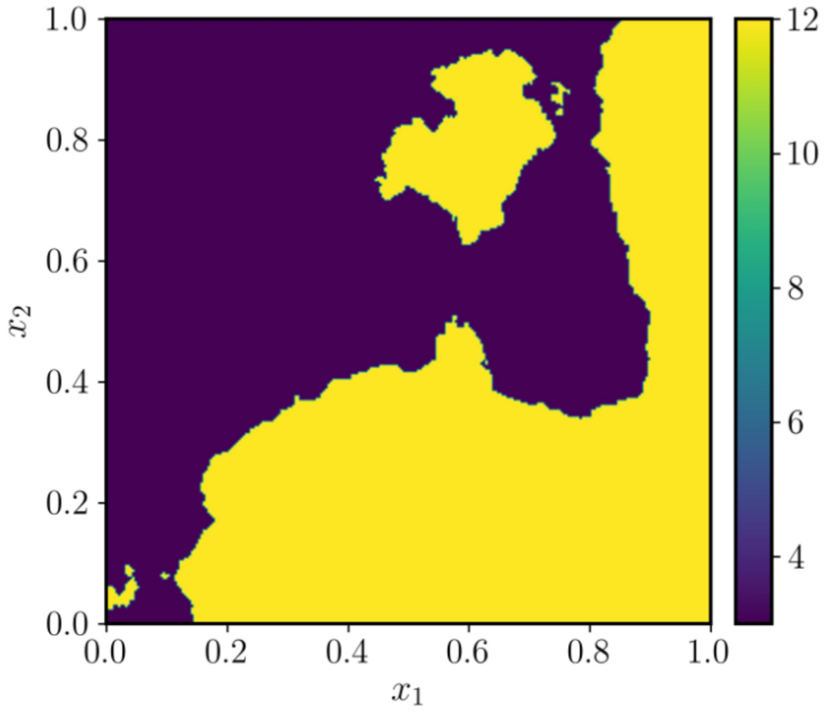
a $\qquad$ u

随机特征方法

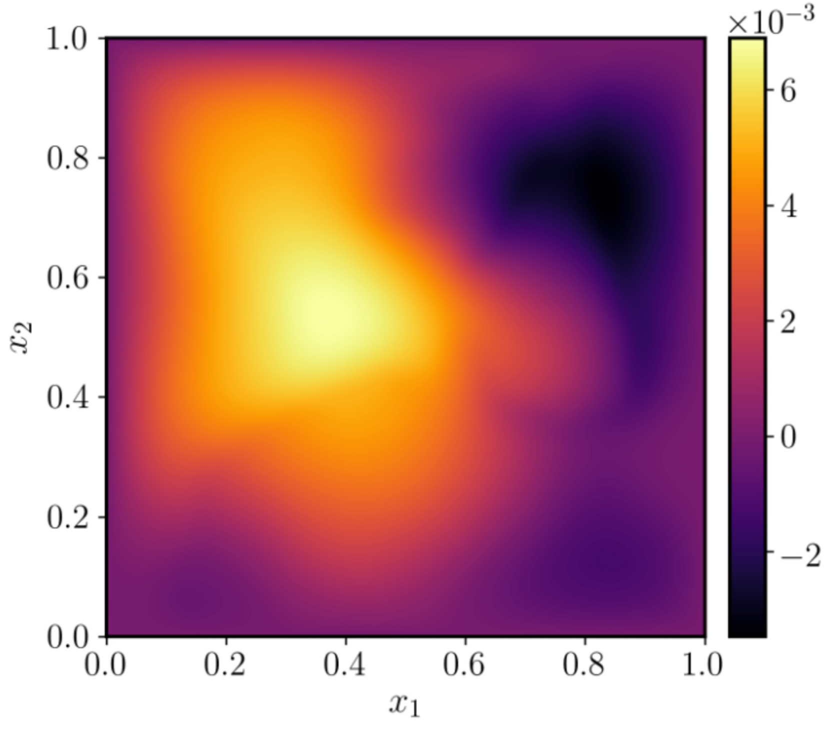
- Fourier随机特征

$$\varphi(a,\omega):\mathcal{A}\times\Theta\rightarrow\mathcal{U}$$

$$\varphi(a,\omega)=\sigma(\mathcal{F}^{-1}\text{filter}(\mathcal{F}a\,\mathcal{F}\omega))$$



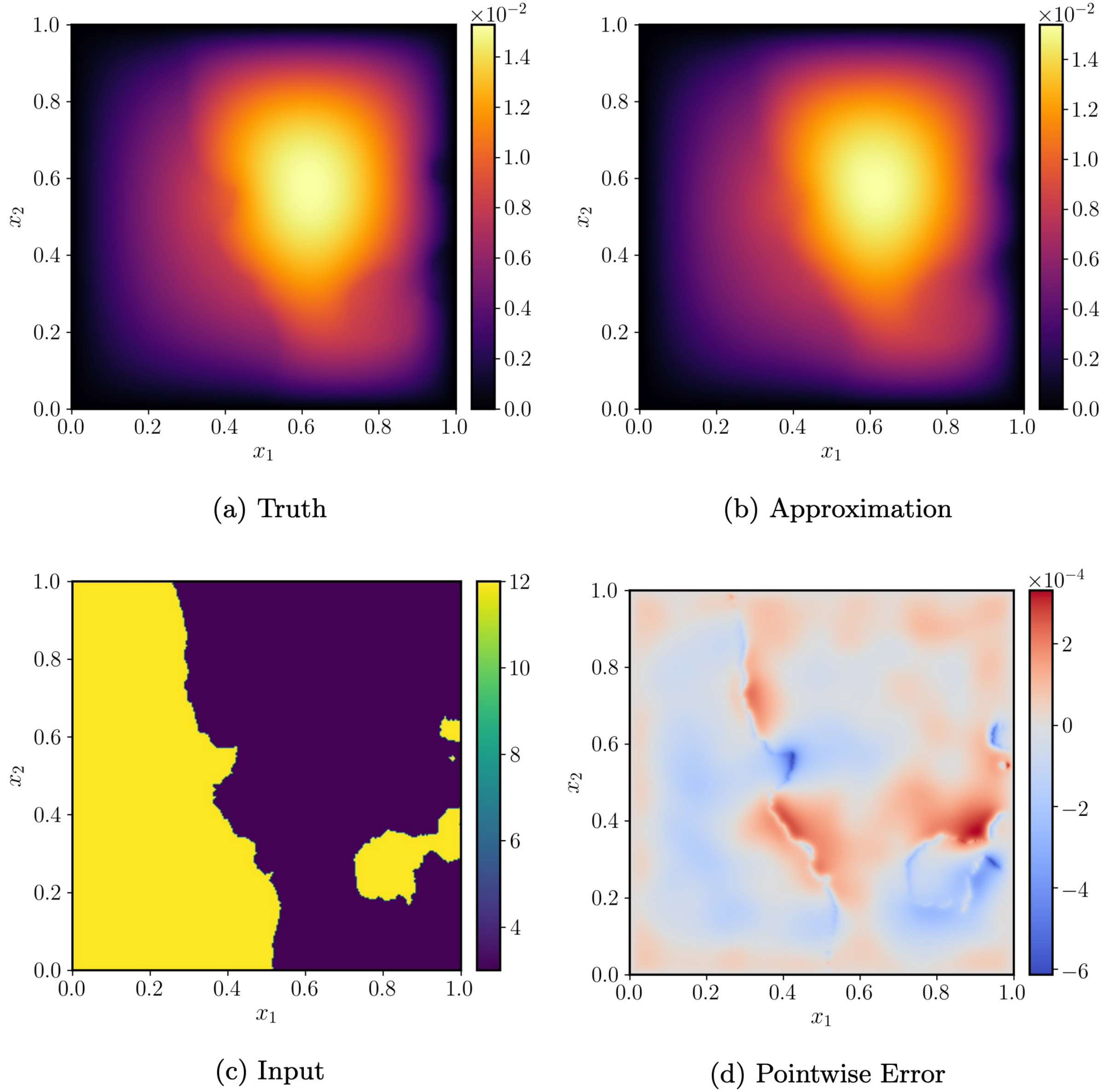
a



$\varphi(a,\omega)$

例子

- Darcy 流方程



高斯回归

- 算子学习

$$
\begin{tikzcd}
\mathcal{A} \arrow[r, "\mathcal{G}^\dagger"] \arrow[d, shift left, "\phi"] & \mathcal{U} \arrow[d, shift left, "\varphi"] \\
\mathbb{R}^n \arrow[u, shift left, "\psi"] \arrow[r, "f^\dagger"] & \mathbb{R}^m \arrow[u, shift left, "\chi"]
\end{tikzcd}
$$

训练数据:

$$\phi(a)=\{a(x_1),a(x_2),\cdots,a(x_n)\}\in R^n$$
$$\varphi(u)=\{u(y_1),u(y_2),\cdots,u(y_m)\}\in R^m$$

目标：

$$\mathcal{G}^\dagger=\chi\circ f^\dagger\circ\phi:\mathcal{A}\mapsto\mathcal{U}$$

高斯过程

- 算子学习

$$
\begin{tikzcd}
\mathcal{A} \arrow[rr, "\mathcal{G}^\dagger"] \arrow[d, shift left, "\phi"] & & \mathcal{U} \arrow[d, shift left, "\varphi"] \\
\mathbb{R}^n \arrow[rr, "f^\dagger"] \arrow[u, shift left, "\psi"] & & \mathbb{R}^m \arrow[u, shift left, "\chi"]
\end{tikzcd}
$$

$\hat{f} \coloneqq \mathrm{argmin}\|f\|_{\mathcal{H}_k}$, 满足 $f(\phi(a_i)) = \varphi(u_i)$

$\hat{u} \coloneqq \mathrm{argmin}\|u\|_{\mathcal{H}_k}$ 满足 $\varphi(u) = \hat{f}(\phi(a)) \in R^m$

含参数偏微分方程问题

- 传统方法(有限体积、有限元方法)

  - 准确但耗时

- 基于高斯回归的方法（随机特征方法、高斯过程）

  - 模型训练、预测均很高效但不精确
  - 需要大量数据训练
  - 超参选取
  - 预测仅限于内插（分布内数据）

参考文献

➢ 高斯过程求解偏微分方程

高斯回归：Zhang, Xiong, Kang Zhu Song, Ming Wan Lu, and X. Liu. "Meshless methods based on collocation with radial basis functions." *Computational mechanics* 26 (2000): 333-343.

高斯回归： Chen, Yifan, Bamdad Hosseini, Houman Owhadi, and Andrew M. Stuart. "Solving and learning nonlinear PDEs with Gaussian processes." *Journal of Computational Physics* 447 (2021): 110668.

随机特征方法：Chen, Jingrun, Xurong Chi, and Zhouwang Yang. "Bridging traditional and machine learning-based algorithms for solving PDEs: the random feature method." *J Mach Learn* 1 (2022): 268-98.

随机特征方法(算子学习)： Nelsen, Nicholas H., and Andrew M. Stuart. "The random feature model for input-output maps between banach spaces." SIAM Journal on Scientific Computing 43, no. 5 (2021): A3212-A3243.

高斯回归(算子学习)： Batlle, Pau, Matthieu Darcy, Bamdad Hosseini, and Houman Owhadi. "Kernel methods are competitive for operator learning." Journal of Computational Physics 496 (2024): 112549.