arXiv:2601.05029v1 [math.OC] 8 Jan 2026

一类贪心型算法在配置优化问题中的随机收敛

E. NIELEN，O. TSE

摘要。贪心采样方法（GSM）被广泛用于构造配置优化问题（COP）的近似解，其中损失泛函在紧域中的有限点构型上进行最小化。尽管贪心型算法在实践中效果显著，但其确定性收敛分析通常受到较多限制，且难以验证。

我们提出了一个随机框架，在该框架中，贪心型方法被表述为配置空间上的连续时间马尔可夫过程。基于这一观点，在误差泛函和转移核满足温和结构假设的条件下，可以对期望意义下和概率意义下的收敛性进行分析。对于全局误差泛函，我们推导出了显式的收敛速率，包括对数衰减、幂律衰减和指数衰减，其具体形式取决于一个抽象的改进条件。

作为教学示例，我们研究了一维分段线性插值中的随机贪心采样，并证明了 L^1-插值误差对于 C^2-函数的指数收敛性。受这一分析启发，我们引入了随机多面体划分方法（R-PDM），这是经典多胞体划分方法的一种随机变体，并通过数值实验展示了其有效性和方差缩减效果。

1. 引言

sec: 引言

贪心采样方法（GSM）用于许多应用领域，包括函数逼近（[2, 9]）、约化基方法（[6, 8, 13, 19, 20, 22, 23, 24, 25, 27,30]）、插值（[1, 4, 5, 17, 26]）等（[18, 26, 31]）。其中许多问题可以重新表述为配置优化问题（COP），并可使用贪心采样方法来逼近其解。在确定性设定下，贪心方法的收敛保证可能难以获得，且可能受到限制。本文提出随机贪心采样方法，并推导出其高概率收敛结果。

配置优化问题。贪心采样方法常用于逼近配置优化问题的解。在配置优化问题中，目标是在给定紧集 $P \subset \mathbb{R}^d$内，以点的构型为自变量，最小化损失函数。在此设定下，一个包含 $n \in \mathbb{N}$ 个来自P 的点的构型 η 是

$$\Omega_n \coloneqq \left\{\eta = (p_1, \dots, p_n, \phi, \phi, \dots) \mid p_i \in P\right\} \subset P^{\mathbb{N}}$$

其中， ϕ 是所谓的墓地状态。序列空间随后由不交并给出

$$\Omega \coloneqq \bigsqcup_{n\geq 0} \Omega_n.$$

日期：2026年1月9日。

对于每个 $n \in \mathbb{N}$，我们进一步定义映射

$$\Omega_n \ni \eta \mapsto \Lambda_n(\eta) = \{p_1, \dots, p_n\} \in \Gamma(P), \qquad \eta = (p_1, \dots, p_n, \phi, \phi, \dots),$$

其中， $\Gamma(P) = \{A \subset P : \#A < \infty\}$ 表示 P的有限子集族 .

配置优化问题表述如下:

问题：COP **问题 1.1.** 设 $P \subset \mathbb{R}^d$ 为一个紧集， $\ell : \Gamma(P) \to [0, +\infty)$ 为给定的损失函数。对于固定的 $n \in \mathbb{N}$，我们的目标是找到

(COP) problem:copproblem:cop

$$\gamma \in \arg\min_{\eta \in \Omega_n} \mathscr{L}(\eta), \qquad \mathscr{L} = \ell \circ \Lambda.$$

注 1.2. 我们注意到，COP 的定义不同于 [21]。这里给出的设定使我们能够追踪各点被添加的顺序，从而使下文考虑的随机过程成为马尔可夫过程。

我们研究的初始动机源于模型降阶中的约化基方法（RBM）。在 RBM 中，目标是逼近解流形 $\mathcal{M} = \{u(p) \in \mathcal{V} : p \in P\}$，其中 $u(p)$ 是由希尔伯特空间 $\mathcal{V}$中的参数 p 所控制的偏微分方程的解。给定一个构型 $\eta \in \Omega_n$，约化基是集合 $\{u(p) : p \in \Lambda_n(\eta)\}$，其张成空间$V_\eta$ 是一个逼近解流形 $\mathcal{M}$的线性空间。损失函数可以定义为

$$\mathscr{L}(\eta) = \max_{q \in P} \left\| u(q) - \mathrm{Proj}_{V_\eta} u(q) \right\|_{\mathcal{V}}^2,$$

其中 Proj_{V_η} 是到线性空间 V_η上的投影算子。其他示例还包括经验插值法 [1], 、最优实验设计 [28], 以及用于回归的主动学习 [32]。

由于通常无法为 COP 找到精确解，因此采用贪心方法来近似求解。这些方法通过迭代构造构型 $\eta \in \Omega$。一般而言，方法从 $\eta^1 = (p_1, \phi, \dots) \in \Omega$开始，其中 $p_1 \in P$是任意选取的。在后续步骤中，通过选择一个新的点 $p_{j+1} \in P$ 并设置$\eta^{j+1} = (p_1, \dots, p_j, p_{j+1}, \phi, \dots)$来更新构型 η^j 。点 p_{j+1} 的选择准则取决于具体的贪心方法。对于形如 $\mathscr{L}(\eta) = \max_{p \in P} J(p, \eta)$的损失函数，经典做法是针对给定的误差函数 $J : P \times \Omega \to [0, +\infty)$，选择满足下式的 p_{j+1}

eq: p 更新 p 更新式： p 更新

$$p_{j+1} \in \arg\max_{q \in P} J(q, \eta^{j-1}). \tag{1.1}$$

换言之，我们选择 P中误差值最高的点，希望该点能够估计出最佳的更新。由于通常无法精确计算 (1.1)，因此这种贪心策略往往会被以下方法取代

(1.2) 弱贪心式：弱贪心式：弱贪心式

$$p_{j+1} \in \arg\max_{q \in S} J(q, \eta^{j-1}),$$

其中， $S \subset P$是一个离散采样集。我们将这一策略称为弱贪心采样。

文献综述。还存在其他贪心策略。在 [29], 中，作者从多个起始点执行梯度下降，以近似全局最大值对应的参数（1.1）。许多替代方法都围绕弱贪心采样中的样本集展开。弱贪婪策略的质量取决于样本集 S对参数集 P的代表性。在实际实现中，样本集 S 的样本大小 $|S|$受到维数灾难的影响。因此，存在许多旨在克服这些可扩展性问题的替代方法，这些方法通常会在每个更新步骤之后考虑不同的样本集 S^j 。例如， [12], 会从 S中移除无关样本，并添加新的、可能相关的样本。在 [11], 中，还会根据验证集中的误差对样本集进行自适应扩充。在 [27], 中，弱贪婪算法在若干个更小的、彼此不相交的训练集上执行；而在 [14], 中，作者使用连续最大化方法构建代理训练集。在 [21], 中，作者引入了多胞体划分方法（PDM）。在 PDM 中，参数集 P 被划分为多个多胞体，样本集由这些多胞体的重心组成。在本文中，我们介绍了该算法的随机化版本（R-PDM）。

在约化基方法的背景下，贪心方法的收敛性通常基于 Kolmogorov n-宽度 [3, 7]进行研究。Kolmogorov n-宽度被定义为能够使解流形 $\mathcal{M}$ 在上确界范数下的逼近误差最小的n维线性空间。在 [3], 中，作者表明，如果能够保证所选参数的误差至少达到精确最大误差的$\gamma \in (0, 1]$ 倍，那么 Kolmogorov n-宽度的多项式或指数衰减分别意味着弱贪婪算法的多项式或指数衰减。由于这一最大误差通常是未知的，因此这一假设很难得到保证。因此，在 [7], 中，他们研究了为以高概率保证该假设成立所需的样本大小$|S|$ ，但仅限于解映射关于参数具有解析性的情形。这些假设不仅难以保证，而且与 Kolmogorov n-宽度的比较也可能具有限制性。解流形中的离群点可能主导收敛速率，并可能造成过于保守的判断。

文献 [16], 提出了另一种视角，将收敛估计与度量熵数进行比较。度量熵数表示用半径为 2^n 的球覆盖一个紧集所需的最小半径。这种视角比经典的 Kolmogorov n-宽度比较得出更紧的界，但仍依赖相同的假设。本文研究了另一种比较方式：将贪心方法建模为随机过程，并推导出概率收敛结果。

论文概述。随机贪心方法的类别比确定性贪心方法更广，因为将核 λ 设为狄拉克测度即可得到确定性方法。将贪心方法重新表述为随机过程有三方面益处：

(1) 概率收敛和期望收敛的陈述比最大误差收敛的陈述限制更少。(2) 可以使用随机方法证明概率收敛和期望收敛。(3) 随机过程的视角使我们能够更全面地看待贪心方法；尤其是，引入核 λ 可以带来新的贪心型算法。

为了将贪心方法建模为随机过程，我们考虑一个随时间变化的构型 η ——构型 $\eta = (\eta_1, \dots, \eta_n, \phi, \dots)$ 跃迁为构型

$$\eta \oplus y = (y, \eta_1, \dots, \eta_n, \phi, \dots) \quad \text{with rate } \lambda(\eta, dy).$$

在第2节中，我们详细介绍该 $\oplus$-算子以及该过程的生成元的精确定义。

在第3节中，我们证明，在一组最少的假设下（参见假设 3.1），对于一大类核 λ，函数 $J : P \times \Omega \to [0, +\infty)$ 无需速率条件即可收敛。由于这些函数还依赖于 P 中的点，我们称它们为局部函数。

对于全局函数 $\mathscr{G} : \Omega \to [0, +\infty)$，例如 (COP) 中的损失函数 $\mathscr{L}$，我们在第4节中证明，在对 $\mathscr{G}$ 和核 λ （参见假设 4.1）施加更严格的假设时，可获得对数、 多项式乃至指数收敛速率（参见定理 4.3）

在第5节中，我们将分段线性插值作为 COP 的一个教学示例。我们证明，使用满足性质 5.8 的转移核的随机贪心方法，一般 C^2-函数的 L^1 插值误差呈指数收敛。

在第 6 节中，我们介绍随机多面体划分方法（R-PDM），并通过数值实验展示了该方法在三个教学示例中的收敛结果。我们还通过数值实验研究了 R-PDM 相较于均匀核在降低方差方面的表现，突显了 R-PDM 的另一项优势。

本文的主要贡献如下:

(1) 我们将贪心方法表述为随机过程。(2) 在一系列不同的假设下，我们推导了这些方法的收敛结果，其中既包括带收敛速率的结果，也包括不带收敛速率的结果。(3) 我们介绍了 R-PDM，并对具有非平凡特性的教学示例进行了分析和数值研究。

2. 构型构造作为随机过程sec: st

机过程

在本节中，我们将贪心型算法建模为随机过程，并在一定假设下证明其适定性。该随机过程对构型中粒子的选择进行建模，即其状态空间为 $\Omega := \bigsqcup_{n \geq 0} \Omega_n$，其中

$$\Omega_n := \left\{ \eta = (\eta_1, \dots, \eta_n, \phi, \dots) \,|\, \eta_i \in P \right\},$$

and ϕ 表示一个墓地状态。此外，我们考虑 Ω上的以下度量

定义：度量

定义 2.1。 令 η， $\sigma \in \Omega$。 Ω 上的度量 $\mathfrak{d}$ 定义为

$$\mathfrak{d}(\eta, \sigma) = \sum_{i \in \mathbb{N}} \frac{1}{2^i} \bar{\mathfrak{d}}(\eta_i, \sigma_i),$$

其中

$$\begin{cases} \bar{\mathfrak{d}}(p, q) := |p - q|_2 & \text{for } p, q \in P, \\ \bar{\mathfrak{d}}(p, q) := \operatorname{diam}(P) & \text{for } p \in P, q = \phi, \text{ or } p = \phi, q \in P, \\ \bar{\mathfrak{d}}(p, q) := 0 & \text{for } p = q = \sigma. \end{cases}$$

然后，我们赋予 Ω 由度量 $\mathfrak{d}$诱导的 Borel σ-代数 $\mathcal{B}_\Omega$。

如前所述，一个构型 η 通过移动其元素并添加一个新点跃迁到另一个构型 σ，可用 $\oplus$-算子表示。

定义 2.2。 算子 $\oplus\colon \Omega \times P \to \Omega$ 定义为

$$\eta \oplus y := (y, \eta_1, \dots, \eta_n, \phi, \dots), \qquad \eta \in \Omega_n,\ y \in P.$$

此外，我们将计数函数 $\mathsf{N}\colon \Omega \to \mathbb{N}$ 定义为

$$\mathsf{N}(\eta) = n, \qquad \eta \in \Omega_n,$$

统计构型中粒子数量 η 它们是 P的元素。[1]

引理 2.3。 度量空间 $(\Omega, \mathfrak{d})$ 是紧的。此外，映射 $\oplus$ 和 N 是连续的，因此是博雷尔可测的。

艾现：compactness omega

该引理的证明见附录 A。

为了刻画随时间发生的可能跃迁，我们现在引入过程的生成元。设 $B_b(\Omega)$ 为 Ω上的有界博雷尔函数集合。过程的生成元 $L\colon B_b(\Omega) \to B_b(\Omega)$ 定义为

$$LF(\eta) := \int_P \big[F(\eta \oplus y) - F(\eta)\big]\, \lambda(\eta, dy). \tag{2.1}$$

方程：生成元 方程：生成元方程：生成元

这里， $F \in B_b(\Omega)$ 是一个观测量， λ 是转移核。在贪婪采样方法的背景下， $F(\eta)$的一个例子可以是某个固定点 $q \in P$处的误差函数 $J(q, \eta)$。给定该生成元，我们假设序列 η 以速率 $\lambda(\eta, dy)$跃迁至 $\eta \oplus y$。 $\lambda(\eta, dy)$的一个简单例子是 P上的均匀测度。

在第 6 节中，我们介绍随机多面体划分方法（R-PDM），并构造一个转移核 λ，使得该过程对应于基于 R-PDM 构造构型。我们始终假设初始条件具有如下形式： $\eta_0 = (\eta_0, \phi, \dots)$，其中 $\eta_0 \in P$。

全文中，我们假设转移核 λ 满足

假设：速率函数

假设 2.4。 转移核 $\lambda\colon \Omega \times \mathcal{B}_{\mathbb{R}^d} \to [0, +\infty)$ 满足

(1) $\lambda(\eta, \cdot) \in \mathcal{P}(P)$, (2) 对任意 $A \in \mathcal{B}_{\mathbb{R}^d}$, 映射 $\Omega \ni \eta \mapsto \lambda(\eta, A)$ 博雷尔可测, (3) 对每个 $\eta \in \Omega$, $\lambda(\eta, P) = 1$。

注 2.5。 我们注意到，这里的结果可以毫无困难地推广到转移核满足 $\sup_{\eta \in \Omega} \lambda(\eta, P) < +\infty$ 的情形。

在假设 2.4 下，我们得到以下存在性结果；只需证明 λ 可导出一个定义良好的转移核 $\kappa\colon \Omega \times \mathcal{B}_\Omega \to [0, +\infty)$, 该结果即可由 [10, §4.2] 得出。为求完备，相关证明见附录 B。

命题 2.6. 设 λ 是满足假设 *2.4* 的转移核。则存在唯一的 Ω 值马尔可夫过程 $(\eta_t)_{t \geq 0}$，其有界生成元 $L\colon B_b(\Omega) \to B_b(\Omega)$ 定义见 (2.1)。

由于 L 是过程 $(\eta_t)_{t\geq 0}$ 的生成元，时间边缘分布 $\mathsf{P}_t = \mathrm{Law}(\eta_t) \in \mathcal{P}(\Omega)$ 满足前向柯尔莫哥洛夫方程

eq: forward kolmogorov eq: forward kolmogorov

$$\langle F, \mathsf{P}_t\rangle - \langle F, \mathsf{P}_s\rangle = \int_s^t \langle LF, \mathsf{P}_r\rangle\, dr, \quad \text{for all } F \in B_b(\Omega), \tag{FKE}$$

其中 $\langle F,\ \mathsf{P}\rangle \coloneqq \int_\Omega F(\eta)\mathsf{P}(d\eta)$。

我们在 $\mathcal{P}(\Omega)$ 上定义全变差范数如下

def: 全变差

定义 2.7。 对于任意 $\mathsf{P},\ \mathsf{Q} \in \mathcal{P}(\Omega)$，全变差范数定义为

$$\|\mathsf{P}-\mathsf{Q}\|_{\mathrm{TV}} \coloneqq \sup\Big\{ |\langle F, \mathsf{P}\rangle - \langle F, \mathsf{Q}\rangle| \,:\, F \in B_b(\Omega),\ \|F\|_\infty \leq 1 \Big\}.$$

由于前向柯尔莫哥洛夫方程（前向柯尔莫哥洛夫方程）对所有 $F \in B_b(\Omega)$ 都成立，我们可以考察特定可观测量 F 的情形，例如，对于某个固定的 $q \in P$，$F \coloneqq J(q,\eta)$。在下一节中，我们将采用这一策略，获得局部函数的收敛结果。

3. 收敛：局部函数

节：收敛

本节证明，误差函数 $J : P\times\Omega \to \mathbb{R}^+$ 在大时间极限下几乎处处收敛到零。我们对局部误差函数作如下假设：

假设 ion: J

假设　定理 3.1. 局部误差函数 $J : P \times \Omega \to [0,+\infty)$ 满足以下条件：

(1)（有界性）存在一个 $c_0 > 0$，使得对所有 $(\eta, p) \in \Omega \times P$，都有 $J(p,\eta) \leq c_0$。

(2)（单调性）对每个 $(\eta, p, y) \in \Omega \times P \times P$，都有

$$J(p, \eta\oplus y) \leq J(p,\eta).$$

(3)（一致性）对每个 $\eta = (\eta_1, \dots, \eta_n, \phi, \dots) \in \Omega_n$，

$$J(\eta_i, \eta) = 0, \qquad i = 1, \dots, n.$$

(4)（正则性）对于任意 $\eta \in \Omega$，映射 $p \mapsto J(p,\eta)$ 是利普希茨连续的，其利普希茨常数为 L_J，且与 η 无关。

注释 3.2。 在许多实际情形中，局部误差函数 $J(p,\cdot)$ 在点的置换下保持不变，即对于某个函数 $_J : P \times \Gamma(P) \to [0,+\infty)$，有 $J(p,\eta) = _J(p, \Lambda(\eta))$。

在定理 3.3 中，我们阐述了本节的主要论断。引理3.5和引理3.6是证明定理 3.3 的重要步骤。

本节的主要论断是：

定理：J 的收敛

定理 3.3。 设 $J : P\times\Omega \to [0,+\infty)$ 是满足假设 *3.1* 的局部误差函数。那么，

$$\lim_{t\to\infty} \int_\Omega |LJ(p,\eta)|\, \mathsf{P}_t(d\eta) = 0 \qquad \textit{for every } p \in P.$$

特别地，如果对于每个 $p \in P$，$\eta \mapsto LJ(p,\eta)$ 都是下半连续的，那么 $(\mathsf{P}_t)_{t\geq 0} \subset \mathcal{P}(\Omega)$ 在窄拓扑下的每个聚点 P_∞ 都满足

$$J(y,\eta) = 0 \quad \textit{for} \quad \lambda(\eta, dy)\mathsf{P}_\infty(d\eta)\textit{-almost every } (y,\eta) \in P \times \Omega.$$

注 3.4. 特别地，如果 $\lambda(\eta,\cdot)$ 与勒贝格测度 $\mathcal{L}$ 等价，即对于每个 $\eta\in\Omega$，$\lambda(\eta,\cdot)\ll\mathcal{L}$ 且 $\lambda(\eta,\cdot)\gg\mathcal{L}$，那么定理 3.3 的结果意味着对于 $\mathcal{L}\otimes\mathsf{P}_\infty$-几乎所有 $(y,\eta)\in P\times\Omega$，都有 $J(y,\eta)=0$。

下述 引理为证明定理 3.3 的第一条陈述提供了基础 .

[1]引理: bounded int积分 ,

Le引理 3.5。 设 $(\mathsf{P}_t)_{t\geq 0}$ 是前向柯尔莫哥洛夫方程的一个解。则

$$\int_0^\infty \langle (LJ(p,\cdot))^-,\mathsf{P}_r\rangle\,dr\leq c_0 \qquad \textit{for every } p\in P.$$

这里，$(LF)^-$ 表示 (LF)的负部，即 $(LF)^-(\eta):=|\min\{0,LF(\eta)\}|$。

证明。由于 $(\mathsf{P}_t)_{t\geq 0}$满足前向柯尔莫哥洛夫方程，因此我们有，对于 $F_p(\eta):=J(p,\eta)$，tha t

$$\begin{aligned}\mathbb{E}[F_p(\eta_t)]-\mathbb{E}[F_p(\eta_0)]&=\int_0^t\int_\Omega LF_p(\eta)\,\mathsf{P}_s(d\eta)\,ds,\\&=\int_0^t\int_\Omega\Big[(LF_p(\eta))^+-(LF_p(\eta))^-\Big]\mathsf{P}_s(d\eta)\,ds,\\&=-\int_0^t\int_\Omega(LF_p(\eta))^-\mathsf{P}_s(d\eta)\,ds.\end{aligned}$$

最后一步由假设 3.1(4)得出。因此，我们得出

$$0\leq\int_0^t\int_\Omega(LF_p(\eta))^-\mathsf{P}_s(d\eta)\,ds\leq\mathbb{E}[F_p(\eta_0)]\leq c_0.$$

令 t 趋于无穷大后，即可得到该结论。 □

引理: 均匀 co 续性

引理3.6。 对于任意 $p\in P$，映射 $t\mapsto\int_\Omega(LJ(p,\eta))^-\mathsf{P}_t(d\eta)$ 一致连续。

证明。 与前面一样，令 $F_p(\eta):=J(p,\eta),\eta\in\Omega$。于是

$$-2c_0\leq L((LJ(p,\eta))^-)\leq 2c_0 \qquad \text{for every } \eta\in\Omega.$$

因此，

$$\left|\int_s^t\langle L((LJ(p,\eta))^-),\mathsf{P}_r\rangle\,dr\right|\leq 2c_0|t-s|,$$

t由此可知，映射 $t\mapsto\int_\Omega(LJ(p,\eta))^-\mathsf{P}_t(d\eta)$ w 可微 ith

$$\left|\frac{d}{dt}\int_\Omega(LJ(p,\eta))^-\mathsf{P}_t(d\eta)\right|\leq 2c_0,$$

从而可以得出，它是一致连续的。 □

由引理 3.5 和 3.6 可得 $\lim_{t\to\infty}\langle(LJ(q,\cdot))^-,\ \mathsf{P}_t\rangle=0$。另一方面， Ω 的紧致性意味着 $\mathcal{P}(\Omega)$ 中的任意概率测度族都是紧的，因此可以断定序列 $(\mathsf{P}_t)_{t\geq 0}\subset\mathcal{P}(\Omega)$ 存在聚点。

现在我们可以证明定理 3.3。

T 的证明 定理 *3.3*。 如前所述，利用引理 3.5 和 3.6，我们可以得出结论： [15]

$$\lim_{t\to\infty}\int_\Omega |LJ(q,\eta)|\,\mathsf{P}_t(d\eta)=0 \qquad \text{for every } p\in P,$$

其中我们用到了这一事实：对于每个 $\eta\in\Omega$，都有 $(LJ(q,\eta))^+=0$。

至于第二部分，我们考虑任意聚点 $\mathsf{P}_\infty\in\mathcal{P}(\Omega)$，以及一个子序列$(\mathsf{P}_{t_n})_{n\geq 1}$，其中 $t_n\to\infty$，当 $n\to\infty$ 时，使得 $\mathsf{P}_{t_n}\rightharpoonup\mathsf{P}_\infty$。由于假设 $\eta\mapsto LJ(q,\eta)$ 是下半连续的，我们得出

$$\int_\Omega (LJ(q,\eta))^-\,\mathsf{P}_\infty(d\eta)\leq\liminf_{n\to\infty}\int_\Omega (LJ(q,\eta))^-\,\mathsf{P}_{t_n}(d\eta)=0.$$

根据假设 3.1(4)，我们进而推得

$$J(p,\eta)=J(p,\eta\oplus y)\qquad \text{for } \lambda(\eta,dy)\mathsf{P}_\infty(d\eta)\text{-almost every } (y,\eta)\in P\times\Omega.$$

特别地，对于 $p=y$，假设 3.1(2)给出

$$0=J(y,\eta\oplus y)=J(y,\eta)\qquad \text{for } \lambda(\eta,dy)\mathsf{P}_\infty(d\eta)\text{-almost every } (y,\eta)\in P\times\Omega,$$

至此证明结束。 □

在接下来的章节中，我们将为全局误差函数这一类函数$\mathscr{G}:\Omega\to\mathbb{R}^+$推导收敛结果，并推导收敛速率。

4. 收敛：全局函数

节：平均改进

本节考虑一类全局函数 $\mathscr{G}:\Omega\to[0,+\infty)$。我们用以下假设替换假设 3.1

假设：平均误差

假设 4.1. 令 $\mathscr{G}:\Omega\to[0,+\infty)$，则 $\mathscr{G}$ 满足：

(1)（有界性）条目| 有界性存在一个 $c_0\in[0,+\infty)$，使得对于所有 $\eta\in\Omega$，都有 $\mathscr{G}(\eta)\leq c_0$。(2)（单调性）条目：单调性 对于每个 $y\in P$ 和 $\eta\in\Omega$，都有

$$\mathscr{G}(\eta\oplus y)\leq\mathscr{G}(\eta).$$

条目：饱和性

(3) (饱和性质) 对任意 $\eta\in\Omega$,

$$\mathscr{G}(\eta)=\mathscr{G}(\eta\oplus y)\ \ \lambda(\eta,dy)\text{-almost every } y\in P \quad\text{implies}\ \ \mathscr{G}(\eta)=0.$$

item: imp进因子

(3')（改进因子）存在一个 $\gamma\in(0,1)$，$\delta>0$、$\beta\in[0,1]$，并且对于每个 $\eta\in\Omega$，都存在一个集合 $B(\eta)\subset P$，满足 $\lambda(\eta,B(\eta))\geq\delta$，使得

$$\int_{B(\eta)}\big(\mathscr{G}(\eta)-\mathscr{G}(\eta\oplus y)\big)\,\lambda(\eta,dy)\geq\frac{\gamma\delta}{\mathsf{N}^\beta(\eta)}\mathscr{G}(\eta).$$

我们考虑条件 (3) 或 (3')。注意，(3') 蕴含 (3)。

全局误差函数这一类包括局部误差函数的平均函数，即,

$$\mathscr{G}(\eta)=\int_P J(q,\eta)\,dq\qquad \text{for some local function } J.$$

它还包括常见的损失函数

$$\mathscr{L}(\eta) = \sup_{q \in P} J(q, \eta),$$

以及前一种情况：对于某个任意但固定的 $q \in P$， $\mathscr{G}(\eta) = J(q, \eta)$。

定理 3.3 可作相应调整，以得到关于全局误差函数的类似结果，如下述引理所示。

定理: 平均 g收敛 gence

定理 4.2. 设 $\mathscr{G}:\Omega \to [0, +\infty)$ 是满足假设 *4.1* (1)–(3) 的全局误差函数。那么，

$$\lim_{t \to \infty} \int_{\Omega} |L\mathscr{G}(\eta)| \, \mathsf{P}_t(d\eta) = 0.$$

特别地，如果 $\mathscr{G}$ 是下半连续的，那么 $(\mathsf{P}_t)_{t \geq 0} \subset \mathcal{P}(\Omega)$ 在窄拓扑下的每个聚点 P_∞ 都满足

$$\mathscr{G}(\eta) = 0 \qquad \textit{for } \mathsf{P}_\infty\textit{-almost every } \eta \in \Omega.$$

证明. 这个引理的证明与定理 3.3 关于 $F(\eta) = \mathscr{G}(\eta)$ 的证明类似。

类似于定理 3.3 的证明，我们可以得出

$$\lim_{t \to \infty} \int_{\Omega} (L\mathscr{G}(\eta))^{-} \mathsf{P}_t(d\eta) = 0.$$

对于一个聚点 $\mathsf{P}_\infty \in \mathcal{P}(\Omega)$ 和一个子序列 $(\mathsf{P}_{t_n})_{n \geq 1}$，其中 $t_n \to \infty$ 是 $n \to \infty$，且满足 $\mathsf{P}_{t_n} \rightharpoonup \mathsf{P}_\infty$，我们得到

$$\int_{\Omega} (L\mathscr{G}(\eta))^{-} \mathsf{P}_\infty(d\eta) \leq \liminf_{n \to \infty} \int_{\Omega} (L\mathscr{G}(\eta))^{-} \mathsf{P}_{t_n}(d\eta) = 0.$$

因此，

$$\int_{\Omega} \int_{P} \big[\mathscr{G}(\eta) - \mathscr{G}(\eta \oplus y)\big] \lambda(\eta, dy) \, \mathsf{P}_\infty(d\eta) = 0,$$

即，对于 $\lambda \otimes \mathsf{P}_\infty$-几乎处处的 $(y, \eta) \in P \times \Omega$， $\mathscr{G}(\eta) = \mathscr{G}(\eta \oplus y)$。根据 $\mathscr{G}$ 的饱和性质（参见假设 4.1 (3)），我们便可得出该断言成立。□

接下来，我们在 $\mathscr{G}$ 上满足改进因子条件的前提下，给出一个比定理 3.3 和 4.2 更强的结果（参见假设 4.1(3')）；具体而言，(1) 我们得到了明确的收敛速率，且 (2) 无需再要求 $\mathscr{G}$ 下半连续，即可断言 $\mathscr{G}(\eta_t) \approx 0$ 对于时刻 $t \gg 1$ 成立。

thm: 收敛 g速率平均 ge

定理 4.3。 令 $(\eta_t)_{t \geq 0}$ 为由带有转移核 $\lambda:\Omega \times P \to [0, +\infty)$ 的 (2.1) 生成的过程，并假设其满足假设 *2.4*。此外，令 $\mathscr{G}:\Omega \to [0, +\infty)$ 为一个全局误差函数，对某些 $\gamma \in (0, 1)$、$\delta > 0$、$\beta \in [0, 1]$ 满足假设 *4.1*。那么，对每个 $\varepsilon > 0$，都存在一个常数 $c_{\mathscr{G}} > 0$，与 ε 无关，使得

$$\mathbb{P}\big(\mathscr{G}(\eta_t) > \varepsilon\big) \leq \frac{c_{\mathscr{G}}}{\varepsilon} \mathbb{E}[\mathscr{G}(\eta_0)] \, \theta_\beta(t) \qquad \textit{for } t \geq 1,$$

with

$$\theta_\beta(t) = \begin{cases} e^{-\gamma\delta t} & \text{for } \beta = 0, \\ t^{1-\frac{1}{\beta}} & \text{for } \beta \in (0,1), \\ \frac{1}{\log(1+t)} & \text{for } \beta = 1. \end{cases}$$

特别地，当 $t \to \infty$ 时，$\mathscr{G}(\eta_t)$ 依概率收敛到 0 。

证明。 根据马尔可夫不等式，我们有

$$\mathbb{P}\big(\mathscr{G}(\eta_t) > \varepsilon\big) \le \frac{1}{\varepsilon}\mathbb{E}[\mathscr{G}(\eta_t)].$$

因此，我们寻找 $E_t := \mathbb{E}[\mathscr{G}(\eta_t)]$ 的一个上界。对于 $\beta = 0$，我们有 th at

$$-L\mathscr{G}(\eta) = \int_P \big[\mathscr{G}(\eta) - \mathscr{G}(\eta \oplus y)\big]\lambda(\eta, dy) \ge \gamma \int_{B(\eta)} \mathscr{G}(\eta)\,\lambda(\eta, dy) \ge \gamma\delta\mathscr{G}(\eta).$$

因此，

$$\frac{d}{dt}E_t \le -\gamma\delta E_t.$$

因此，根据格朗沃尔不等式，我们得出

$$E_t \le E_0 e^{-\gamma\delta t}.$$

对于 $\beta \in (0,1]$，我们有

$$\int_P \big[\mathscr{G}(\eta) - \mathscr{G}(\eta\oplus y)\big]\lambda(\eta, dy) \ge \int_{B(\eta)} [\mathscr{G}(\eta) - \mathscr{G}(\eta\oplus y)]\lambda(\eta, dy) \ge \frac{\gamma\delta}{\mathsf{N}^\beta(\eta)}\mathscr{G}(\eta).$$

令 $\mu := \gamma\delta$，则对于任意 $K \in [1, +\infty)$，有

$$\begin{aligned} \frac{d}{dt}E_t &\le -\mu\mathbb{E}\left[\frac{\mathscr{G}(\eta_t)}{\mathsf{N}^\beta(\eta_t)}\right] = -\mu\mathbb{E}\left[\frac{\mathscr{G}(\eta_t)}{\mathsf{N}^\beta(\eta_t)}\mathbb{1}_{\{\mathsf{N}(\eta_t)>K\}}\right] - \mu\mathbb{E}\left[\frac{\mathscr{G}(\eta_t)}{\mathsf{N}^\beta(\eta_t)}\mathbb{1}_{\{\mathsf{N}(\eta_t)\le K\}}\right], \\ &\le -\frac{\mu}{K^\beta}\mathbb{E}\left[\mathscr{G}(\eta_t)\mathbb{1}_{\{\mathsf{N}(\eta_t)\le K\}}\right] = -\frac{\mu}{K^\beta}\mathbb{E}\left[\mathscr{G}(\eta_t)\right] + \frac{\mu}{K^\beta}\mathbb{E}\left[\mathscr{G}(\eta_t)\mathbb{1}_{\{\mathsf{N}(\eta_t)>K\}}\right], \\ &\le -\frac{\mu}{K^\beta}E_t + \frac{c_0\mu(1+t)}{K^{1+\beta}} =: g_t(K, E_t). \end{aligned}$$

在最后一步中，我们使用了假设 4.1(1) 以及以下事实:

$$\mathbb{E}\left[\mathbb{1}_{\{\mathsf{N}(\eta_t)>K\}}\right] = \mathbb{P}\big(\mathsf{N}(\eta_t) > K\big) \le \frac{1+t}{K}.$$

现在我们来确定使 $K \mapsto g_t(K, E_t)$ 取得最小值的条件。一个驻点由下式给出

$$K_\circ = \frac{1+\beta}{\beta}\frac{c_0}{E_t}(1+t) \in [1, \infty).$$

由于，$\partial_K^2 g_t(K_\circ,\ E_t) = \beta\mu E_t / K_\circ^{2+\beta} > 0$，驻点 $K_\circ$ 是一个极小值点。于是，我们有 e

$$\frac{d}{dt}E_t \le g_t(K_\circ, E_t) = -\frac{\alpha_\beta}{(1+t)^\beta}E_t^{1+\beta}, \qquad \alpha_\beta := \frac{\mu c_0^{-\beta}}{1+\beta}\left(\frac{\beta}{1+\beta}\right)^\beta.$$

求解 E_t 的微分不等式，得到

$$E_t \le E_0 \left(1 + \alpha_\beta E_0^\beta \frac{\beta}{1-\beta}\Big((1+t)^{1-\beta} - 1\Big)\right)^{-1/\beta} \le E_0\, \mathcal{O}\big(t^{1-\frac{1}{\beta}}\big)|_{t\to\infty}.$$

最后，对于 $\beta = 1$，我们推导得出

$$E_t \le E_0\Big(1 + \alpha_1 E_0 \log(1+t)\Big)^{-1} = E_0\, \mathcal{O}\Big(\big(\log(1+t)\big)^{-1}\Big)|_{t\to\infty},$$

由此完成证明。 □

5. 教学示例：一维插值

sec: 插值

COP 的一个例子是分段线性插值。本节的主要结果（定理 5.9）指出，在转移核 λ 满足某些假设的条件下，C^2-函数的分段线性插值的 L_1 误差呈指数收敛。

在陈述主要结果之前，我们先定义分段线性插值。分段线性插值函数取决于节点集 $x_0 < x_1 < \ldots < x_{n+1}$，其中有 $x_0 = a$ and$x_{n+1} = b$。我们使用随机贪心方法在参数集 $P = [a,b]$中寻找插值节点 $\{x_1, \ldots, x_n\}$。请注意，点 $x_0 = a$ 和 $x_{n+1} = b$ 不属于算法选出的节点。通常，随机贪心方法不会生成有序列表$\eta = (x_1, \ldots, x_n, \phi, \ldots)$。因此，我们定义以下映射来排列 $\eta \in \Omega$的元素。

定义：序映射

定义 5.1. 序映射 $\mathfrak{S}\colon \Omega \to \Omega$ 定义为

$$\mathfrak{S}(\eta) \coloneqq (\eta_{\sigma(1)}, \ldots, \eta_{\sigma(\mathsf{N}(\eta))}, \phi, \ldots),$$

suc使得 $x_i = \eta_{\sigma(i)}$ 对 $i \in \{1, \ldots, \mathsf{N}(\eta)\}$ 满足 $x_i \le x_{i+1}$，对所有 $i \in \{0, \ldots, \mathsf{N}(\eta)$均成立 $\}$.

$\mathfrak{S}$ 的一个重要性质见下述引理，其证明为完整起见见附录 A。

引理：序映射

引理 5.2。 序映射 $\mathfrak{S}\colon \Omega \to \Omega$ 是连续的。

借助一个序 环 η，现在可以基于 η定义函数的线性近似 .

定义 5.3 （线性插值）。 设 $f\colon [a,b] \to \mathbb{R}$ 为一个函数。对于任意 $\eta \in \Omega_n$，相对于 η 的 f 的分段线性近似定义为

linear approximation Leq: linear approximation L

$$\mathfrak{I}_\eta[f](x) = f(x_k) + \frac{f(x_{k+1}) - f(x_k)}{x_{k+1} - x_k}(x - x_k) \qquad \text{for } x \in [x_k, x_{k+1}], \tag{5.1}$$

其中 $\mathfrak{S}(\eta) = (x_1, \ldots, x_n, \phi, \ldots)$。

注记 5.4. 注意，我们可以将 $\mathfrak{I}_\eta[f]$ 改写为

$$\mathfrak{I}_\eta[f](x) = f(a)\sigma_a(x) + \sum_{i=1}^{n} f(x_i)\sigma_{x_i}(x) + f(b)\sigma_b(x), \qquad \mathfrak{S}(\eta) = (x_1, \ldots, x_n, \phi, \ldots),$$

with

$$\sigma_{x_i}(x) = \begin{cases} \frac{x-x_{i-1}}{x_i-x_{i-1}} & \text{for } x \in [x_{i-1}, x_i], \\ \frac{x_{i+1}-x}{x_{i+1}-x_i} & \text{for } x \in [x_i, x_{i+1}], \\ 0 & \text{otherwise.} \end{cases}$$

注意，由于 $\mathsf{N}(\eta) \to \infty$，该表达式可求和，因为

$$\sum_{i=1}^{n} f(x_i)\sigma_{x_i}(x) \le \|f\|_{\sup} \sum_{i=1}^{n} \sigma_{x_i}(x) \le \|f\|_{\sup}.$$

注记 5.5。 在线性近似（5.1）的定义中，我们假设插值节点严格递增。我们注意到，定义 5.1 并未要求点$\{x_{\sigma(1)}, \dots, x_{N(\eta)}\}$也满足这一条件。然而，对于许多随机贪心方法，重复抽取相同点的概率为 0。

我们考虑以下局部误差函数

差 插值
差插值 q: latione 逐点误差 r interpolation

$$J(x,\eta) = |\mathfrak{T}_\eta[f](x) - f(x)|, \qquad x \in P := [a,b]. \tag{5.2}$$

我们关注以下全局误差函数的收敛性。

q: error interpolationeq: error interpolation

$$\mathscr{G}(\eta) = \int_P J(x,\eta)\,dx = \|\mathfrak{T}_\eta[f] - f\|_{L^1}. \tag{5.3}$$

对于这个全局误差函数，以下定理成立

定理 5.6. *5.3* 中定义的全局误差函数 $\mathscr{G}:\Omega \to [0,+\infty)$ 是连续的。

证明. 注意，只需证明 $J(x,\eta)$ 是连续的。设 $\eta \in \Omega$，令 $(\eta^k)_{k\in\mathbb{N}}$是 Ω 中的一个序列，满足当 $k \to \infty$ 时 $\mathfrak{d}(\eta^k,\eta) \to 0$ 。那么存在一个 $K \in \mathbb{N}$，使得对于所有 $k \ge K$都有$\mathsf{N}(\eta) = \mathsf{N}(\eta^k)$。令 $\mathfrak{S}(\eta) = (x_1, \dots, x_n, \dots)$，并令对于所有$k \ge K$都有 $\mathfrak{S}(\eta^k) = (x_1^k, \dots, x_n^k, \dots)$，再定义一个函数 $T^k:[a,b] \to [a,b]$，使其具有以下性质：对于所有 $i \in \{0,\dots,n\}$和 $k \ge K$，都有 $T^k([x_i^k, x_{i+1}^k]) = [x_i, x_{i+1}]$。该函数为

$$T^k(x) = \frac{x_{i+1}^k - x}{x_{i+1}^k - x_i^k} x_i + \frac{x - x_i^k}{x_{i+1}^k - x_i^k} x_{i+1} \quad \text{for } x \in [x_i^k, x_{i+1}^k].$$

成立的是

$$\mathfrak{T}_{\eta^k}[f](x) - \mathfrak{T}_\eta[f](x) = \underbrace{\mathfrak{T}_{\eta^k}[f](x) - \mathfrak{T}_{\eta^k}[f]\big((T^k)^{-1}(x)\big)}_{(A)} + \underbrace{\mathfrak{T}_{\eta^k}[f]\big((T^k)^{-1}(x)\big) - \mathfrak{T}_\eta[f](x)}_{(B)}.$$

对于 (A)，有

$$(A) \le \|\mathfrak{T}_{\eta^k}[f]\|_{\sup} |x - (T^k)^{-1}(x)| \le \|f\|_{\sup} |x - (T^k)^{-1}(x)|.$$

注意到

$$(T^k)^{-1}(x) = \frac{x_{i+1} - x}{x_{i+1} - x_i} x_i^k + \frac{x - x_i}{x_{i+1} - x_i} x_{i+1}^k \quad \text{for } x \in [x_i, x_{i+1}],$$

于是得到估计

$$\begin{aligned}|x-(T^k)^{-1}(x)| &= \left|\frac{(x_{i+1}-x_i)x-(x_{i+1}-x)x_i^k-(x-x_i)x_{i+1}^k}{x_{i+1}-x_i}\right| \\ &= \left|\frac{(x-x_i)(x_{i+1}-x_{i+1}^k)+(x_{i+1}-x)(x_i-x_i^k)}{x_{i+1}-x_i}\right| \\ &\le \max\{|x_{i+1}-x_{i+1}^k|,|x_i-x_i^k|\}.\end{aligned}$$

因此，我们得出结论

$$(A)\le \|f\|_{\sup}\max\{|x_{i+1}-x_{i+1}^k|,|x_i-x_i^k|\}.$$

此外，根据定义，我们有

$$\mathfrak{T}_{\eta^k}[f](x)\overset{(*)}{=}\mathfrak{T}_{\eta}[f\circ(T^k)^{-1}](T^k(x)).$$

因此，我们有

$$\begin{aligned}(B)&\overset{(*)}{=}|\mathfrak{T}_{\eta}[f\circ(T^k)^{-1}](x)-\mathfrak{T}_{\eta}[f](x)| \\ &=|\mathfrak{T}_{\eta}[f\circ(T^k)^{-1}-f](x)| \\ &=\left|\frac{x-x_i}{x_{i+1}-x_i}(f\circ(T^k)^{-1}-f)(x_{i+1})+\frac{x_{i+1}-x}{x_{i+1}-x_i}(f\circ(T^k)^{-1}-f)(x_i)\right|, \\ &=\left|\frac{x-x_i}{x_{i+1}-x_i}(f(x_{i+1}^k)-f(x_{i+1})+\frac{x_{i+1}-x}{x_{i+1}-x_i}(f(x_i^k)-f(x_i))\right|, \\ &\le \|f\|_{\sup}\max\{|x_{i+1}-x_{i+1}^k|,|x_i-x_i^k|\}.\end{aligned}$$

综上，我们得到

$$\|J(\cdot,\eta^k)-J(\cdot,\eta)\|_{\sup}=\|\mathfrak{T}_{\eta^k}[f]-\mathfrak{T}_{\eta}[f]\|_{\sup}\le 2\|f\|_{\sup}\max_{i=1,\dots,n}|x_i^k-x_i|.$$

我们得出结论，该表达式趋于零，因为 $\mathfrak{S}$ 连续。因此，根据控制收敛定理， $J(x,\cdot)$ 和 $\mathcal{G}(\cdot)$ 也都是连续的。□

为了表述我们的主要结果，我们需要定义集合 $B_\mu(\eta)$。

定义：B(eta)**定义5.7**。 设 $\mu\in(0,1/2)$、 $\eta\in\Omega_n$、 $x_0=a$， $x_{n+1}=b$ 和 $\mathfrak{S}(\eta)=(x_1,\dots,x_n,\phi,\dots)$。对于每个 $k\in\{0,\dots,n\}$，我们令 $I_k^\mu:=(x_k+\mu(x_{k+1}-x_k),x_{k+1}-\mu(x_{k+1}-x_k))$。于是，

eq: B(eta) B(eta)el: B(eta)

$$B_\mu(\eta):=\left\{x\in[a,b]:x\in I_k^\mu,\ k\in\{0,\dots,n\}\right\}. \tag{5.4}$$

我们对转移核作出以下假设，稍后将在第 6 节证明，均匀采样和 R-PDM 满足该假设。

假设：插值跃迁 核

性质 5.8. 令 $\mu\in(0,1/2)$. 转移核 $\lambda:\Omega\times\mathcal{B}_{\mathbb{R}^d}\to[0,+\infty)$ 满足

$$\lambda(\eta,B_\mu(\eta))\ge 1-2\mu\quad\text{for all }\eta\in\Omega.$$

下面给出关于收敛的主要结果。

定理：插值 C2 **定理 5.9.** 设 $f \in C^2([a,b])$ 满足 $c_f := \|f''\|_{\sup}$。此外，设 $\mu \in (0,1/2)$，并令 (η_t) 为由满足性质 *5.8* 的转移核 λ生成的过程。则存在常数$m_{\mathscr{G}} > 0$，使得对于任意 $\varepsilon > 0$ 和 $\alpha > c_f$,

$$\mathbb{P}\big(\mathscr{G}(\eta_t) > \varepsilon\big) \le \frac{m_{\mathscr{G}}}{\varepsilon} e^{-\gamma\delta t} \qquad \textit{for } t \ge 1,$$

其中 $\delta := 1 - 2\mu$、 $\gamma := \mu\frac{\alpha - c_f}{\alpha + c_f}$。

为了证明此定理，我们首先证明，对于强凸 C^2-函数，假设 4.1(1)–(3') 成立。由此可推得，收敛结果对于 C^2-函数也成立。其思路如下：由于 $f \in C^2([a,b])$，令 $h_\alpha := \alpha|x|^2/2$，则函数

$$x \mapsto f_\alpha(x) := f(x) + h_\alpha(x) \qquad \text{is strongly convex for } \alpha > c_f.$$

因此，$f = f_\alpha - h_\alpha$ 其中 f_α 和 h_α 均为强凸函数。这样，我们得到

$$\begin{aligned}\|\mathfrak{T}_\eta[f] - f\|_{L_1} &= \|(\mathfrak{T}_\eta[f_\alpha] - f_\alpha) - (\mathfrak{T}_\eta[h_\alpha] - h_\alpha)\|_{L_1} \\ &\le \|\mathfrak{T}_\eta[f_\alpha] - f_\alpha\|_{L_1} + \|\mathfrak{T}_\eta[h_\alpha] - h_\alpha\|_{L_1}.\end{aligned}$$

我们对强凸函数有如下结果 $f \in C^2([a,b])$。

引理：满足假设 **引理 5.10.** 设 $f \in C^2([a,b])$ 是强 m-凸函数，即，

$$f((1-r)x + ry) \le (1-r)f(x) + rf(y) - \frac{m}{2}r(1-r)|x-y|^2 \qquad \textit{for every } r \in [0,1].$$

设 $\mathscr{G}:\Omega \to [0,+\infty)$ 由 (5.3) 给出。设 $\mu \in (0,1/2)$，且转移核 λ 满足性质 *5.8*。则满足假设 *4.1*(1)–*(3')*。

为证明此引理，我们必须检验假设 4.1(1)–(3') 是否成立。我们分步骤证明这一点。首先，我们给出一个引理，说明局部误差函数

$$J : [a,b] \times \Omega \to [0,+\infty), \quad (p,\eta) \mapsto J(p,\eta) := |\mathfrak{T}_\eta[f](p) - f(p)|,$$

满足假设 3.1。

引理：插值 p误差递减 g **引理 5.11.** 设 $f \in C^2[a,b]$ 是一个强 m-凸函数，且在 $[a,b]$ 上满足 $0 < m \le f'' \le M$。则有以下结论：

*(1) (*一致性*)* 对任意 $\eta \in \Omega_n$，若满足 $\mathfrak{S}(\eta) = (x_1, \cdots, x_n, \phi, \ldots)$，则有 $J(x_i, \eta) = 0$。

(2) 对任意满足 $\mathfrak{S}(\eta) = (x_1, \cdots, x_n, \phi, \ldots)$ 的 $\eta \in \Omega$，对于 $p \in (x_k, x_{k+1})$，有

$$0 \le \frac{m}{2}(x_{k+1} - p)(p - x_k) \le \mathfrak{T}_\eta[f](p) - f(p) \le \frac{M}{2}(x_{k+1} - p)(p - x_k).$$

*(3) (*单调性*)* 设 $y \in (x_k, x_{k+1})$。则对于任意 $p \in [a,b]$,

$$J(p,\eta) - J(p,\eta \oplus y) \ge \frac{m}{M}\left[\frac{x_{k+1} - y}{x_{k+1} - p}\mathbf{1}_{(x_k,y)}(p) + \frac{y - x_k}{p - x_k}\mathbf{1}_{(y,x_{k+1})}(p)\right] J(p,\eta) \ge 0.$$

特别地，J 是一个满足假设 *3.1* 的局部误差函数。

证明。令 $\eta \in \Omega_n$ 和 $p \in [a,b]$。此外，令 $\mathfrak{S}(\eta) = (x_1, \dots, x_n, \phi, \dots)$、$x_0 = a$ 和 $x_{n+1} = b$。那么，对于某个 $k \in \{0, \dots, n\}$，有 $p \in [x_k, x_{k+1}]$。

首先注意到，对于任意 $k \in \{0, \dots, n\}$，若 $p = x_k$，则根据插值算子的定义可得 $J(p,\eta) = 0$，从而得到式 (1)。

对于式 (2)，我们利用 f 的强 m-凸性推得

$$f(p) \le \frac{x_{k+1}-p}{x_{k+1}-x_k} f(x_k) + \frac{p-x_k}{x_{k+1}-x_k} f(x_{k+1}) - \frac{m}{2}(x_{k+1}-p)(p-x_k),$$

由此我们得到

$$\begin{aligned}\mathfrak{I}_\eta[f](p) - f(p) &= f(x_k) + \frac{f(x_{k+1})-f(x_k)}{x_{k+1}-x_k}(p-x_k) - f(p) \\ &\ge \frac{m}{2}(x_{k+1}-p)(p-x_k).\end{aligned}$$

对于上界，我们利用泰勒公式得到

$$\mathfrak{I}_\eta[f](p) - f(p) \le \frac{M}{2}(x_{k+1}-p)(p-x_k).$$

综合以上结果，可得对每个 $p \in (x_k, x_{k+1})$，该断言均成立。

现在证明 (3)：假设 $y \in I_k := (x_k, x_{k+1})$ 对某个 $k \in \{0, \dots, n\}$成立。

情况 *1*：$p \notin I_k$。在这种情况下，我们直接得到 $J(p, \eta\oplus y) = J(p,\eta)$，因为误差的变化只发生在区间 I_k 内。

情况 *2*：$p = y$。由 (1) 可得 $J(p,\eta\oplus y) = 0 \le J(p,\eta)$。

情况 *3*：$p \in I_k$，$p \ne y$。与 (2) 的证明类似，我们得到

$$\mathfrak{I}_\eta[f](p) - \mathfrak{I}_{\eta\oplus y}[f](p) \ge \frac{m}{2}(x_{k+1}-y)(p-x_k) \qquad \text{for } p \in (x_k, y),$$

and

$$\mathfrak{I}_\eta[f](p) - \mathfrak{I}_{\eta\oplus y}[f](p) \ge \frac{m}{2}(x_{k+1}-p)(y-x_k) \qquad \text{for } p \in (y, x_{k+1}).$$

将这些估计合并起来，可得

$$\begin{aligned}J(p,\eta) - J(p,\eta\oplus y) &= \mathfrak{I}_\eta[f](p) - \mathfrak{I}_{\eta\oplus y}[f](p) \\ &\ge \frac{m}{M}\left[\frac{x_{k+1}-y}{x_{k+1}-p}\mathbf{1}_{(x_k,y)}(p) + \frac{y-x_k}{p-x_k}\mathbf{1}_{(y,x_{k+1})}(p)\right]\big(\mathfrak{I}_\eta[f](p) - f(p)\big),\end{aligned}$$

这就是该断言中的第 (3) 点。 □

我们注意到，引理 5.11 直接意味着全局误差函数

$$\mathcal{G}:\Omega \to [0,+\infty), \quad \eta \mapsto \mathcal{G}(\eta) := \int_a^b J(q,\eta)\,dq,$$

满足 $\mathcal{G}(\eta\oplus y) \le \mathcal{G}(\eta)$。

在证明引理 5.10 之前，我们再陈述一个引理，以便得出 $\mathcal{G}$ 满足假设 4.1(3') 。

引理：插值假设 3 **引理 5.12**。 设 $f \in C^2[a,b]$ 是一个具有 $0 < m \le f'' \le M$ 的强 m-凸函数，定义在 $[a,b]$ 上。此外，设 $\mu \in (0,1/2)$、$\delta \coloneqq (1-2\mu) > 0$ 和 $B_\mu(\eta)$ 由 (5.4) 给出。若 $\lambda(\eta, B_\mu(\eta)) \ge \delta$，则

$$\int_{B_\mu(\eta)} \big[\mathscr{G}(\eta) - \mathscr{G}(\eta \oplus y)\big]\, \lambda(\eta, dy) \ge \frac{\delta \mu m}{M} \mathscr{G}(\eta) \qquad \textit{for every } \eta \in \Omega.$$

证明。 由引理 5.11(3)，我们推得

$$\begin{aligned}
\int_{B_\mu(\eta)} \big[\mathscr{G}(\eta) - \mathscr{G}(\eta \oplus y)\big]\, \lambda(\eta, dy) &= \iint_{[a,b]\times B_\mu(\eta)} \big[J(q,\eta) - J(q, \eta\oplus y)\big]\, \lambda(\eta, dy)\, dq \\
&\ge \frac{m}{M} \sum_{k=0}^{n-1} \iint_{[a,b]\times I_k^\mu} \left[\frac{x_{k+1}-y}{x_{k+1}-q} \mathbf{1}_{(q, x_{k+1})}(y) + \frac{y-x_k}{q-x_k} \mathbf{1}_{(x_k, q)}(y) \right] J(q,\eta) \lambda(\eta, dy)\, dq \\
&\ge \mu \frac{m}{M} \sum_{k=0}^{n-1} \lambda(\eta, I_k^\mu) \int_{[a,b]} J(q,\eta)\, dq = \mu(1-2\mu)\frac{m}{M} \mathscr{G}(\eta),
\end{aligned}$$

其中我们运用了富比尼定理交换积分次序，以及这一事实 t

$$\frac{x_{k+1}-y}{x_{k+1}-q} \mathbf{1}_{(q, x_{k+1})}(y) + \frac{y-x_k}{q-x_k} \mathbf{1}_{(x_k, q)}(y) \ge \mu \mathbf{1}_{(x_k, x_{k+1})}(y) \qquad \text{for almost every } y \in I_k^\mu. \qquad \square$$

引理 *5.10* 的证明。 我们证明假设 4.1 中的所有条目均得到满足。

(1) 令 $x_0 = a$ 和 $x_1 = b$。令

$$\mathfrak{I}_0[f](x) = \frac{f(b)-f(a)}{b-a}(x-a) + f(a).$$

那么对于任意 $\eta_0 = (p, \phi, \dots)$，引理 5.11 蕴含

$$\mathscr{G}(\eta_0) \le \int_P |\mathfrak{I}_0[f](x) - f(x)| \le c_0,$$

其中

$$c_0 = (b-a) \cdot \max_{x\in[a,b]} |\mathfrak{I}_0[f](x) - f(x)|.$$

(2) 这是引理 5.11 的推论。

(3') 设 $\gamma = \frac{\mu m}{M} \in (0,1)$、$\beta = 0$、$\delta = 1 - 2\mu > 0$，并对每个 $\eta \in \Omega$，令 $B_\mu(\eta)$ 由 (5.4) 给出。此项由引理 5.12 得出。

$\square$

定理 *5.9* 的证明。 由于 $f \in C^2([a,b])$，令 $h_\alpha \coloneqq \alpha^{\frac{|x|^2}{2}}$，我们有函数

$$x \mapsto f_\alpha(x) \coloneqq f(x) + h_\alpha(x) \qquad \text{is strongly convex for } \alpha > c_f,$$

因为

$$m_{f_\alpha} \coloneqq \alpha - c_f \le f''(x) \le \alpha + c_f \coloneqq M_{f_\alpha},$$

对所有 $x \in [a,b]$ 均成立。此外， h_α 是强凸的，因为对所有 $x \in [a,b]$，$h''_\alpha(x) = \alpha$ 均成立。而且，

$$\begin{aligned}\|\mathfrak{I}_{\eta_t}[f] - f\|_{L_1} &= \|\mathfrak{I}_{\eta_t}[f_\alpha - h_\alpha] - (f_\alpha - h_\alpha)\|_{L_1} \\ &\le \|\mathfrak{I}_{\eta_t}[f_\alpha] - f_\alpha\|_{L_1} + \|\mathfrak{I}_{\eta_t}[h_\alpha] - h_\alpha\|_{L_1}.\end{aligned}$$

设 $\mu \in (0, 1/2)$，并设 $\varepsilon > 0$。我们可以仿照引理 5.10 的证明，并利用定理 4.3 的结果得出

$$\mathbb{P}\left(\|\mathfrak{I}_{\eta_t}[f_\alpha] - f_\alpha\|_{L_1} > \frac{\varepsilon}{2}\right) \le \frac{2}{\varepsilon}\|\mathfrak{I}_{\eta_0}[f_\alpha] - f_\alpha\|_{L_1} e^{-\gamma_{f_\alpha}\delta t},$$

其中 $\gamma_{f_\alpha} := \mu \frac{\alpha - c_f}{\alpha + c_f}$，以及

$$\mathbb{P}\left(\|\mathfrak{I}_{\eta_t}[h_\alpha] - h_\alpha\|_{L_1} > \frac{\varepsilon}{2}\right) \le \frac{2}{\varepsilon}\|\mathfrak{I}_{\eta_0}[h_\alpha] - h_\alpha\|_{L_1} e^{-\gamma_{h_\alpha}\delta t},$$

其中 $\gamma_{h_\alpha} = \mu$。

我们注意到，对每一个 $\alpha > c_f$ 都有 $\gamma_{f_\alpha} \le \gamma_{h_\alpha}$。设 $m_{\mathscr{G}} := 4\max\{\|\mathfrak{I}_{\eta_0}[f_\alpha] - f_\alpha\|_{L_1}, \|\mathfrak{I}_{\eta_0}[h_\alpha] - h_\alpha\|_{L_1}\}$，则有

$$\begin{aligned}\mathbb{P}\left(\|\mathfrak{I}_{\eta_t}[f] - f\|_{L_1} > \varepsilon\right) &= \mathbb{P}\left(\|\mathfrak{I}_{\eta_t}[f_\alpha] - f_\alpha\|_{L_1} > \frac{\varepsilon}{2}\right) + \mathbb{P}\left(\|\mathfrak{I}_{\eta_t}[h_\alpha] - h_\alpha\|_{L_1} > \frac{\varepsilon}{2}\right) \\ &\le \frac{m_{\mathscr{G}}}{\varepsilon} e^{-\gamma_{f_\alpha}\delta t} \quad \text{for } t \ge 1.\end{aligned}$$

至此证明完毕。 □

6. 随机多面体划分方法

sec: rpdm

本文中的若干结果依赖于关于转移核 λ 的一般假设。本节将考察该转移核的具体形式，并证明这些核满足相关假设。具体而言，我们考察与随机化的多胞体划分方法（R-PDM）相对应的转移核，以及均匀转移核。我们首先介绍 R-PDM。

6.1. 随机多面体划分方法。 R-PDM 是一种将超矩形 参数集 P 划分为多个多胞体，并搜索局部误差 J 较大区域的算法。每个构型 $\eta = (\eta_1, \dots, \eta_n, \phi, \dots)$ 对应于一组特定的多胞体 $\mathcal{D}(\eta)$，这些多胞体将参数集 P 划分开来（$\mathcal{D}(\eta)$ 的详细构造见下文）。与 R-PDM 相对应的转移核定义为

(6.1) eq: 速率函数 rpdm eq: 速率函数 rpdmeq: 速率函数 rpdm

$$\lambda_{\text{rpdm}}(\eta, dy) := \sum_{D \in \mathcal{D}(\eta)} \frac{e^{\alpha \fint_D J(q,\eta)\,dq}}{Z} \operatorname{unif}_D(dy),$$

其中 $\alpha > 0$、 Z 是归一化因子， unif_D 是多胞体 D 上的均匀测度。

这个转移核纯属理论构造，因为在实际中，通常无法精确计算 $\fint_D J(q,\eta)\,dq$。因此，在实际应用中，我们通过以下方式近似这个积分

误差近似 approximationeq: 平均误差近似

$$\fint_D J(q,\eta)\,dq \approx J(p_D, \eta), \quad (6.2)$$

其中 $p_D \in D$。作为初始近似，令 $p_D \coloneqq \mathsf{b}_D$，其中 b_D 是D的重心。由于重心可能导致积分 $\int_D J(q,\eta)\,dq$的近似效果较差，因此，当 $J(\mathsf{b}_D,\eta) < \epsilon$ 满足某个预设容差 ϵ时，我们可以用任意一点 p_D 替换 b_D。

备注 6.1：式（6.2）给出的近似（其中 $p_D = \mathsf{b}_D$ ）是使用一个求积点的中点法则。通过使用更多求积点，或使用随机积分器，可以改进该近似，但会牺牲计算效率和可扩展性。

转移核 λ 取决于多胞体划分 $\mathcal{D}(\eta)$。在 R-PDM 中，经过 η 次跃迁后，该划分会细化为一个新的状态 $\eta \oplus y$。这种细化基于一种称为面连接的操作，其定义如下：如 [21] 所示，

定义 6.2。设 $P \subset \mathbb{R}^d$ 为一个多胞体，$p \in P$ 为任意一个点。此外，设 ∂P表示 P的面集合。那么，面连接算子 FL 定义为

$$\mathsf{FL}(p, P) = \left\{\mathrm{Conv}(p \cup F) \;:\; F \in \partial P\right\}.$$

换言之，面连接算子通过将一个点 p 连接到所有面$F \in \partial P$来划分多胞体。多胞体划分 $\mathcal{D}(\eta)$ 以下列方式取决于面连接算子：设 η 是 R-PDM 中具有多胞体划分 $\mathcal{D}(\eta)$的一个状态，并且对于某个 $y \in P$ 有$y \in D \in \mathcal{D}(\eta)$，则 $\eta \mapsto \eta \oplus y$ $\mathcal{D}(\eta \oplus y) = \mathcal{D}(\eta) \setminus D \cup \mathsf{FL}(y, D)$。图 1 展示了二维情况下 R-PDM 前几个步骤的示例。该方法总结如下算法：

算法 1 随机多面体划分方法

rpdm

1：初始化点的数量 n、常数 $\alpha > 0$ 和容差 $\epsilon > 02$：$k \leftarrow 13$：选择 $p \in \mathrm{int}(P)4$：设置 $\eta \coloneqq (p, \phi, \ldots)5$：设置 $\mathcal{D} \coloneqq \{D \in \mathsf{FL}(p, P)\}6$：对所有 $D \in \mathcal{D}$，$p_D = \mathsf{b}_D$ 7：**当** $k < n$**时**8：对所有 $D \in \mathcal{D}$，计算 $J(p_D, \eta)$。9：根据 $e^{\alpha J(p_D,\eta)}$ 加权的概率对 $D \in \mathcal{D}(\eta)$ 进行抽样10：根据 unif_D 对 $y \in D$ 进行抽样11：$\eta \leftarrow \eta \oplus y12$：对所有满足 $J(p_E, \eta) < \epsilon$ 的 $E \in \mathcal{D}(\eta) \setminus D$，在 E 中均匀重采样 $p_E \in E13$：$\mathcal{D} \leftarrow (\mathcal{D} \setminus \{D\}) \cup \mathsf{FL}(y, D)14$：对于所有 $E \in \mathsf{FL}(y, D)$，设置 $p_E = \mathsf{b}_E15$：$k \leftarrow k + 1$

R-PDM 背后的思想是，在 P 误差 J 较高的区域放置更多质量。转移核 λ_{rpdm} 满足假设 2.4。此外，$\lambda_{\mathrm{rpdm}}(\eta, \cdot)$ 与勒贝格测度等价（即 $\lambda_{\mathrm{rpdm}}(\eta, \cdot) \sim \mathcal{L}$），这意味着定理 3.3 和定理 4.2 的收敛结果在 $\mathcal{L}(dy) \otimes P_\infty$-几乎处处成立。对于全局误差函数 $\mathscr{G}$ 满足假设 4.1 的应用，我们通过定理 4.3 建立收敛速率。更具体地说，如果 C^2 函数满足假设 5.8，则定理 5.9 保证其插值收敛。在下面的引理 6.3 中，我们证明 λ_{rpdm}确实满足这一条件。

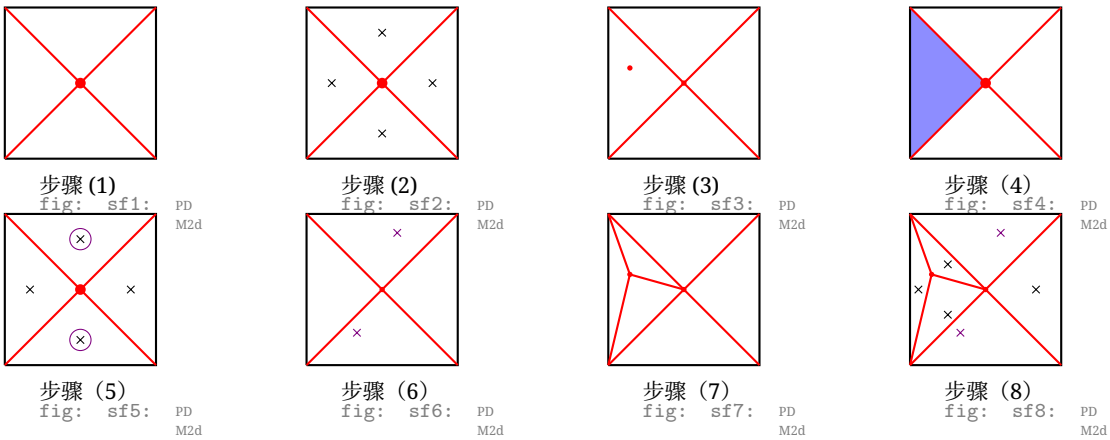


图 1. R-PDM 在2维参数情形下各步骤及通过面连接进行分割域的示意图。（1）首先对参数进行采样，并通过面连接划分 P。（2）计算重心。（3）根据转移核选择一个点。（4）标记包含该点的多胞体。（5）选择其他多胞体中误差函数低于容差的参数。（6）对新点进行采样，以替换这些重心。（7）更新多胞体划分。（8）计算新的重心。图 : PD M2d

然而，还存在其他与勒贝格测度等价并满足相同性质的转移核。一个简单的例子是均匀测度

$$\lambda_{\mathrm{unif}}(\eta, \cdot) \coloneqq \mathrm{unif}_{P}.$$

通过均匀测度选择点在计算效率上高于 R-PDM。然而，计算效率并不是唯一需要考虑的因素。

首先，我们必须指出，定理 4.3 的证明中所用的估计较为粗糙。我们考虑的是对时间一致成立的估计，因此忽略了局部随时间改进可能存在的差异。事实上，在假设 4.1 中，我们假设 $\lambda(\eta, B(\eta)) \geq \delta$ 对某个 $\delta > 0$ 随时间一致成立，由此得到的界不可能对每个转移核都足够紧，因为对于某些 η，$B(\eta)$ 的测度可能远大于 δ。由于 R-PDM 会在误差较大的区域分配更多质量，因此在局部范围内，我们可以预期这一假设对比 λ_{unif} 更大的集合 $B(\eta)$ 成立。

其次，收敛结果以高概率成立，但在实践中，方差可能存在很大差异。较高的方差意味着，尽管误差在期望意义下收敛至零，但算法的单次运行仍可能得到较差的 COP 近似解。用户可能更倾向于选择方差较低的算法，从而对单次运行的收敛结果更有信心。目前，这些额外因素都未体现在理论结果中。我们将给出数值测试以展示它们的重要性，并计算若干贪心型算法的方差。

现在我们证明，λ_{rpdm} 和 λ_{unif} 确实满足性质 5.8，由此可得定理 5.9 的指数收敛结论。

引理：质量 B_eta

引理 6.3。*Le* 令 $\eta \in \Omega$、 $\mu \in (0, 1/2)$和 $B_\mu(\eta)$ 由 (5.4) 给出。则性质 *5.8* 成立 :

$$\lambda_{\text{unif}}(\eta, B_\mu(\eta)) = \lambda_{\text{rpdm}}(\eta, B_\mu(\eta)) = 1 - 2\mu > 0.$$

证明。 我们注意到

$$\lambda_{\text{unif}}(\eta, B_\mu(\eta)) = \sum_{k=0}^{n} \int_{I_k^\mu} \frac{1}{b-a} dy = \sum_{k=0}^{n} \frac{|I_k^\mu|}{b-a} = (1-2\mu) \sum_{k=0}^{n} \frac{x_{k+1} - x_k}{b-a} = 1 - 2\mu > 0.$$

同样，对于 R-PDM，根据构造，对于每个 $D \in \mathcal{D}(\eta)$，恰好存在一个 $k \in \{0, \dots, n\}$ 使得 $I_k^\mu \subset D$。我们也将这个 D 记为 D_k。因此，

$$\begin{aligned}\lambda_{\text{rpdm}}(\eta, B_\mu(\eta)) &= \sum_{k=0}^{n} \int_{I_k^\mu} \sum_{D \in \mathcal{D}(\eta)} \frac{e^{\alpha f_D J(z,\eta)\, dz}}{Z} \text{unif}_D(dy) \\ &= \sum_{k=0}^{n} \frac{e^{\alpha f_{D_k} J(z,\eta)\, dz}}{Z} \frac{|I_k^\mu|}{x_{k+1} - x_k} = (1-2\mu) \sum_{k=0}^{n} \frac{e^{\alpha f_{D_k} J(z,\eta)\, dz}}{Z} = 1 - 2\mu > 0,\end{aligned}$$

由此完成证明。 □

收敛结果适用于 λ_{rpdm} 和 λ_{unif}。在数值实验中，我们还计算了经典弱贪心方法（见 (1.2)）的误差和方差，

备注6.4。弱贪心方法依赖于样本集 $S \subset P$。这一表述的一个结果是，在此应用中不满足假设 4.1 的饱和性质。我们可以通过将假设 4.1 和定理 4.3 中所有的 P 替换为离散集 S，对它们进行重新表述。类似地，误差函数 (5.3) 变为

$$\mathcal{G}_S(\eta) = \sum_{x \in S} J(x, \eta).$$

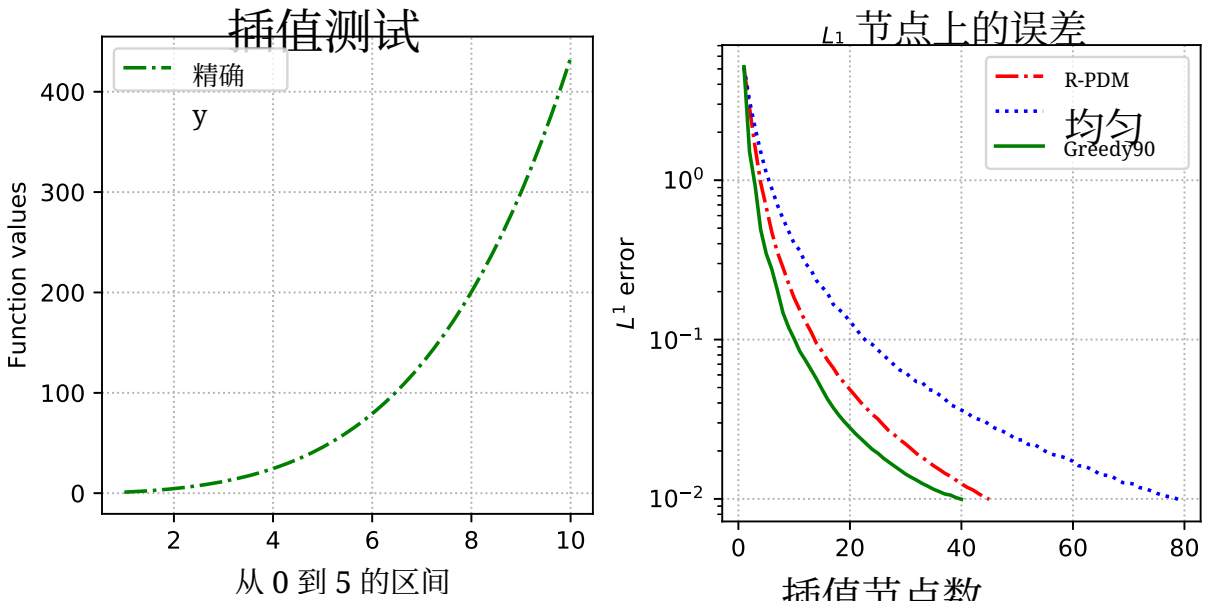
在这一重新表述的框架下，我们可以期待插值框架中的弱贪心方法具有 $\mathcal{G}_S$ 的收敛结果。

6.2. **数值实验**。

6.2.1. 示例*1*。我们采用随机贪心方法，为三个不同的函数计算插值节点。[1]第一个函数 $f : [0, 5] \to [0, +\infty)$定义为 $f(x) = x^2 + \frac{1}{30}x^4$，如图 2a 所示。我们注意到，该函数无穷次可微且强凸，因此根据引理 5.10，假设 4.1 得以满足。对于 R-PDM，我们设定 $\alpha = 500$，容差为 $\epsilon = 0.01$。

对于给定的 η_t，其中有 t 个选定的插值节点，我们可以通过将区间 $[0, 5]$ 离散化为 $L = 500$ 个离散化点 $(y_1, \dots, y_L)$，并在这些离散化点上应用求积规则，来近似误差 $\mathcal{G}(\eta_t)$。为了进行稳健的比较，我们运行算法

[1]本小节中用于生成数值结果的所有代码均可在此处找到：
https://gitlab.tue.nl/s158446/interpolation.git



(A) 精确函数 $f(x) = x^2 + \frac{1}{30}x^4$ 图：递增的精确函数 (B) 插值节点上 L^1-误差的衰减 图：递增的误差衰减

图2. 示例1

$K = 1000$ 次并计算平均误差，即我们计算

ectation interpolationeq: expectation interpolation

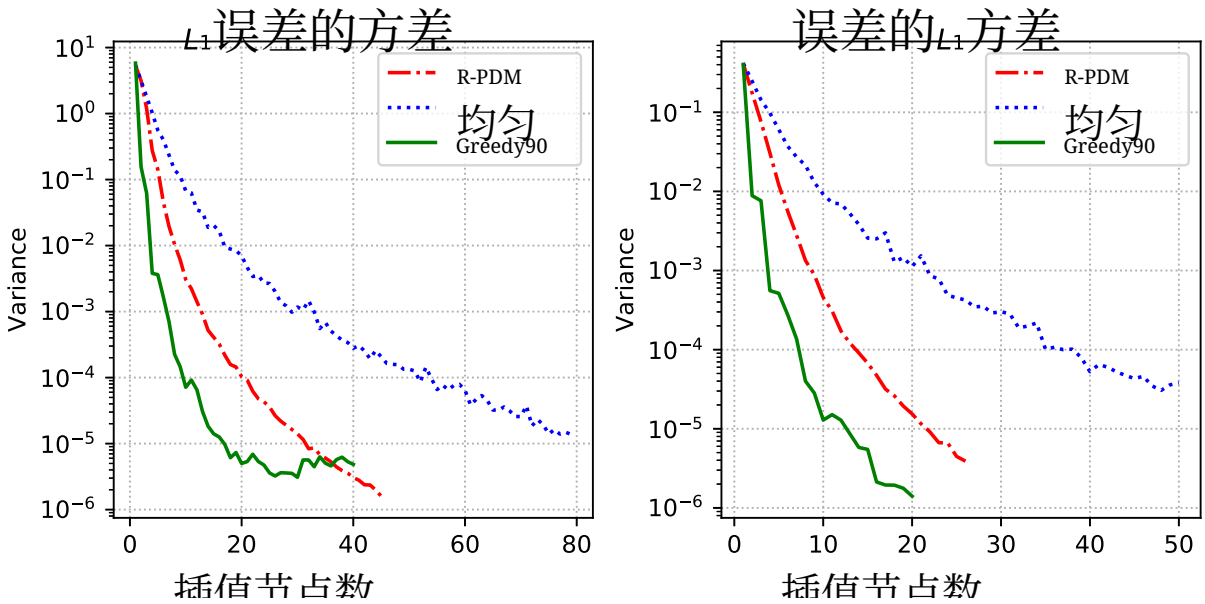
$$E_t := \frac{1}{KL}\sum_{k=1}^{K}\sum_{i=1}^{L} J(y_i, \eta_t^k). \tag{6.3}$$

我们不断添加点，直到 E_t 低于容差 10^{-2}。这意味着，不同算法最终可能得到不同的配置长度。图2b展示了误差随时间步数量的衰减情况。首先可以看到，R-PDM的误差低于均匀采样的误差。因此，R-PDM的选择过程在45个节点处停止，而均匀采样则要到79个节点才停止。我们还运行了弱贪婪算法（见1.2），样本大小设为90。我们注意到，该算法的误差略低于R-PDM。贪心算法的缺点是用户必须事先选择样本大小，而且很难确定该选多大。在这个具体例子中，如果样本大小小于79（即均匀算法选取的节点数），我们不能合理地期望弱贪心方法达到容差0.01。此外，对于弱贪心方法，我们评估了3510个误差估计，而R-PDM只需要进行1034次评估。

我们还比较了两种方法的方差。首先，我们计算 $J(q, \eta_t)$在 t 个时间步上的方差，即计算

variance interpolationeq: variance interpolation

$$V_t := \frac{1}{K}\sum_{k=1}^{K}\left(\frac{1}{L}\sum_{i=1}^{L} J(y_i, \eta_t^k) - E_t\right)^2, \tag{6.4}$$



(A) 每次插值的 L^1-误差方差
示例1的节点

fig: 总方差递增

(B) 每个插值节点数对应的 L^1-误差方差
例 2 的节点

fig: 总方差凸性

图3. 比较示例1和例 2 的 L^1-误差方差

其中 E_t 由 (6.3) 给出。该结果见图3a。起初，贪心方法的方差低于 R-PDM 和均匀采样，但对于最终配置，R-PDM 的方差低于弱贪心采样的方差。尽管选取的插值点数量更多，均匀采样得到的最终配置的方差仍高于 R-PDM 的方差。

我们还计算了最终构型 η_t（即达到 10^{-2} 的 L^1-误差时得到的构型）在离散化点 $(y_1, \dots, y_L)$ 处的平均逐点误差 $J(y_i, \eta_t)$。我们再次对 $K = 1000$ 次运行的误差取平均，即计算

数值逐点误差插值 方程: 数值逐点误差插值方程: 数值逐点误差插值 ation

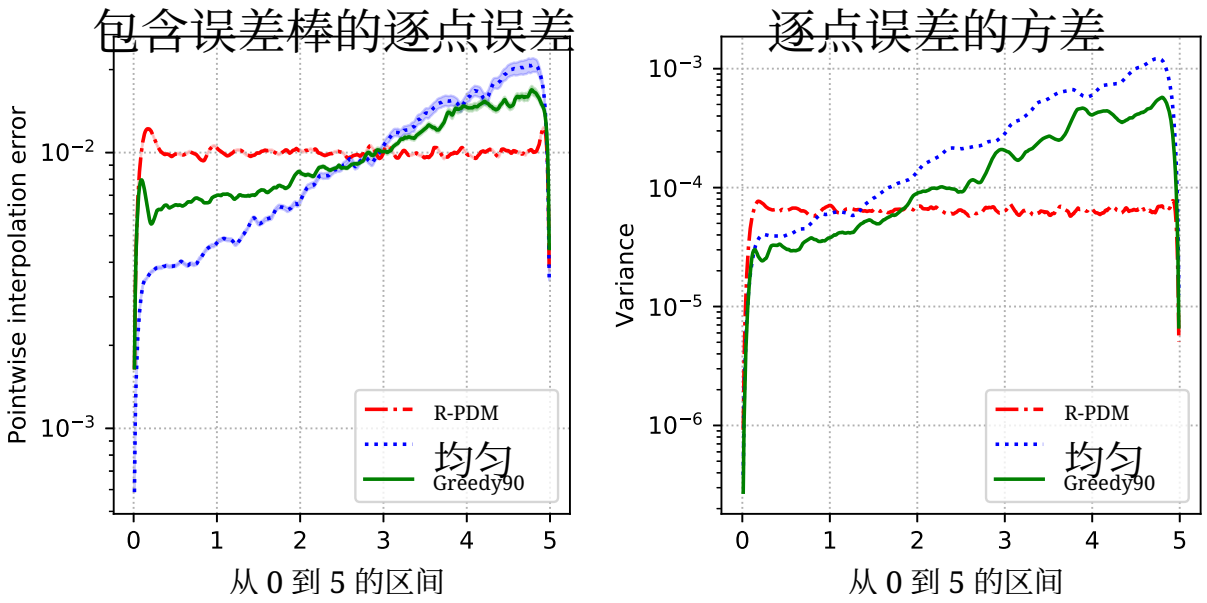
$$E_t(y_i) := \frac{1}{K} \sum_{k=1}^{K} J(y_i, \eta_t^k).$$

图4a展示了这些平均逐点误差，包括误差棒。这些误差棒由两种方法的方差确定，即我们计算了

variance interpolationeq: pointwise variance interpolation

$$V_t(y_i) := \frac{1}{K} \sum_{k=1}^{K} \left(J(y_i, \eta_t^k) - E_t(y_i) \right)^2. \tag{6.6}$$

对于所有方法而言，边界附近的误差都较低，因为根据定义，分段逼近在边界点处始终是精确的。我们注意到，R-PDM 的逐点误差始终约为 10^{-2}（靠近边界处除外）。均匀采样和弱贪心方法在区间左侧的逐点误差较低，而在右侧的数值较高。我们注意到，总体而言，R-PDM 的误差棒似乎小于均匀



（A）离散化点处的误差及误差条 (B) 离散化点的方差

fig: 平均误差递增 fig: 方差递增

图4. 示例1的逐点误差和方差

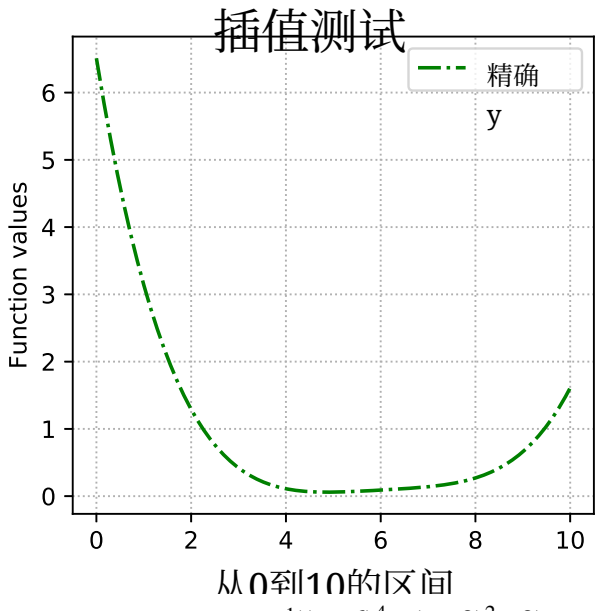
采样和弱贪心方法的误差棒。为清晰起见，区间 $[0,5]$ 中的方差显示在图4b中；我们注意到，均匀采样和弱贪心方法的上确界方差比 R-PDM 的上确界方差大一个数量级。此外，R-PDM 的方差保持在 10^{-4} 附近，而另外两种方法的方差则呈振荡变化。

6.2.2. 例 *2*。 在第二个实验中，我们考虑函数（见图 5a）

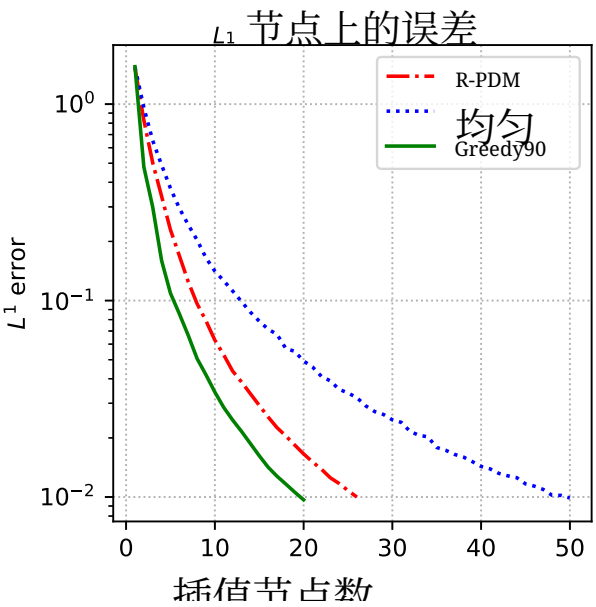
$$f(x)=\frac{1}{200}\left((x-6)^4+(x-2)^2+2\right),\qquad x\in[0,10].$$

我们注意到，该函数同样无穷次可微且强凸，但不再是递增函数。对于 R-PDM，我们设定 $\alpha=500$ 和 $\epsilon=0.01$，并再次分别运行两种随机方法和弱贪心方法 1000 次，选择插值点，直至达到 0.01 的容差。我们再次将区间 $[0,10]$ 离散为 500 个点$(y_1,\dots,y_L)$。由 (6.3) 给出的误差 E_t 显示于图 5b。我们再次注意到，均匀采样方法需要选择多得多的插值节点才能达到相同的误差容差。我们使用 90 个样本运行弱贪心方法，这再次得到最低的 L^1-误差。我们注意到，R-PDM 的误差仅略高一些，但在构造过程中，R-PDM 对逐点误差 (5.3) 进行了 350 次计算，而弱贪心方法则需要进行 1710 次计算。预计在高维问题中，计算样本数量的这一差异会进一步增大（见 [21]）。

时间步方差（见（6.4））如图3b所示。值得注意的是，这一次弱贪心方法的方差最低。弱贪心方法和 R-PDM 的方差均比均匀采样的方差低一个数量级以上



(A) 精确函数 $f(x) = \frac{1}{200}((x-6)^4+(x-2)^2+2)$



(B) 节点上的 L^1-误差衰减

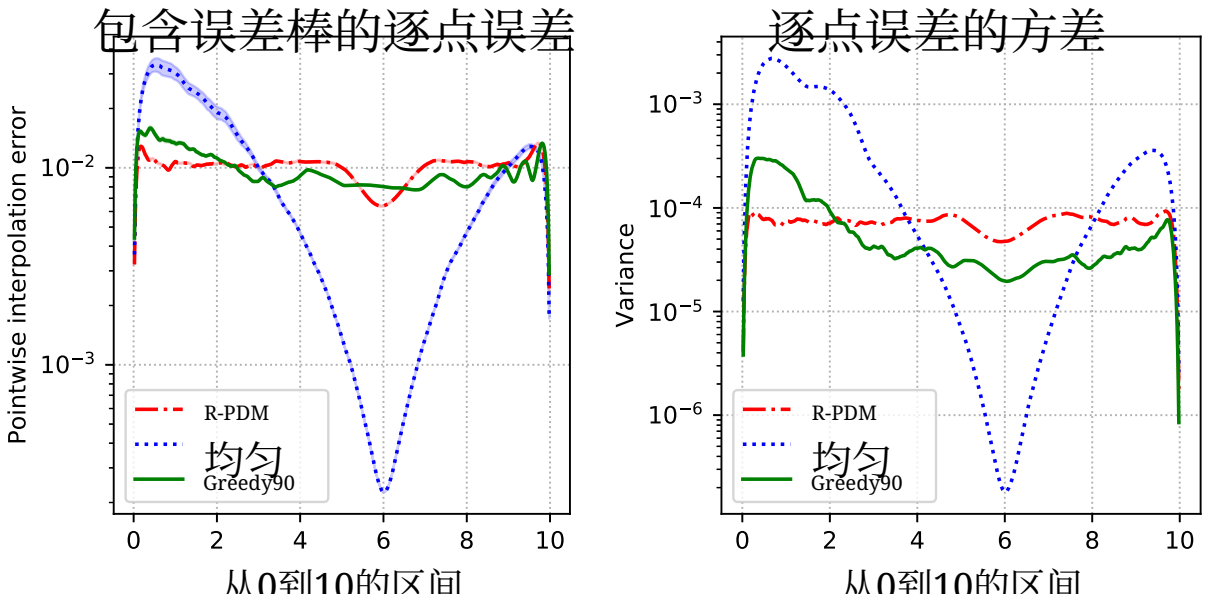
图：凸精确函数 图：凸误差衰减

图 5. 例 2

在最终配置中，尽管均匀采样选定的插值节点数量更多。

最后，图 6a 展示了 500 个离散化点处的逐点误差 (6.5) 及误差棒。这些误差棒同样由方差确定（为便于说明，方差展示在图 6b 中）。在此例中，R-PDM 和弱贪心采样的误差相近。与这两种方法相比，在区间 $[3, 9]$ 内，均匀采样的平均误差较低，而在其他所有位置都较高。从图 5a 可以看出，该区间也对应于函数最平坦的部分，因此也是可以用线性函数合理逼近的部分。由于 R-PDM 和弱贪心采样在其选择过程中考虑了逼近误差，因此与均匀采样相比，它们在该区域选取的插值节点更少。这解释了方差上的差异。除边界附近外，R-PDM 和弱贪心采样的逐点误差方差都保持在 10^{-4} 左右。均匀采样的方差则表现出更大的波动。对于均匀采样，某些区域的方差可能很高，而另一些区域则很低。

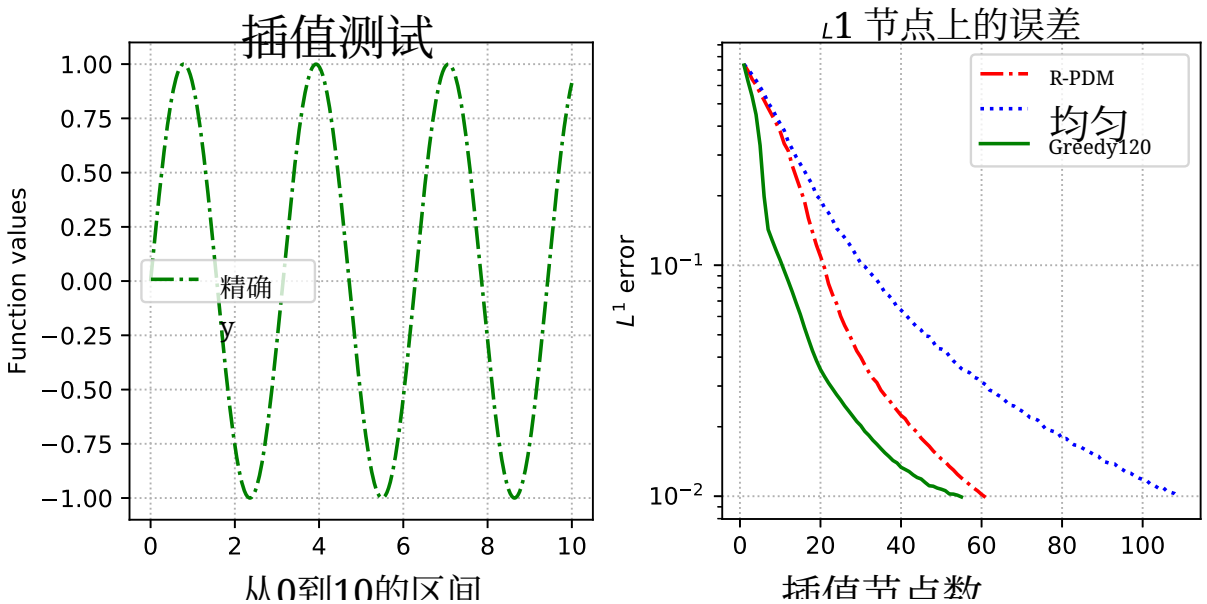
6.2.3. 例 *3.* 在此例中，我们考虑区间 $[0, 10]$ 上的函数 $f(x) = \sin(2x)$（见图 7a）。值得注意的是，该函数仍然无穷次可微，但不再是强凸函数。根据定理 5.9，我们仍然预期此例中会收敛。对于 R-PDM，我们设定 $\alpha = 500$ 和 $\varepsilon = 0.01$，并将每种方法运行 1000 次，选择插值点，直至达到 0.01 的容差。区间 $[0, 10]$ 被离散为 500 个点 $(y_1, \ldots, y_L)$。



（A）离散化点处的误差及误差条 （B）离散化点处的方差

Fig: 凸平均误差 Fig: 凸方差

图 6. 例 2 的逐点误差和方差



（A）精确函数 $f(x) = \sin(2x)$ （B）节点上的 L^1-误差衰减

Fig: 正弦精确函数 Fig: 正弦误差衰减

图 7. 例 3

图 7b 显示了由 (6.3) 给出的平均 L^1-误差。弱贪心方法中使用 120 个样本。我们再次注意到，均匀采样的 L^1-误差收敛最慢

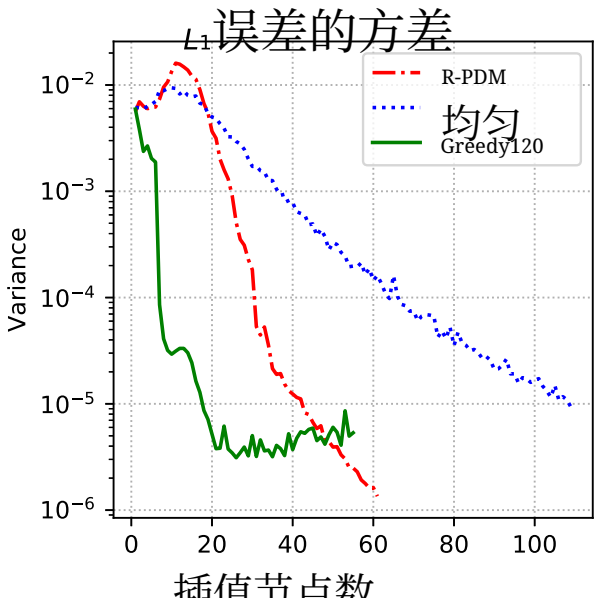
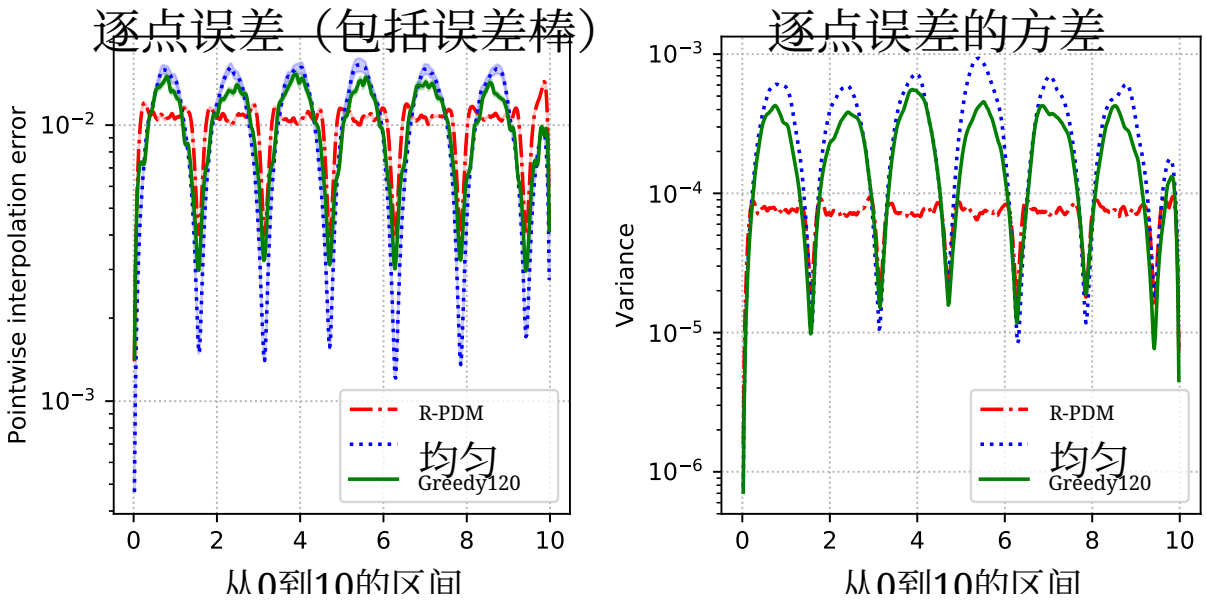


图8. 示例 L^1-每插值节点数的误差方差 fig: 3总方差-sin

对于弱贪心方法，速度最快；与之前一样，弱贪心方法和 R-PDM 之间的差异并不大。需要指出的是，我们对贪心方法评估了 6480 个样本，而对 R-PDM 仅评估了 1890 个样本。 L^1-误差的方差见图 8。弱贪心方法的方差起初最低，但在选定约 20 个插值节点后开始振荡。对于 R-PDM 和均匀采样，方差起初有所增加；但 R-PDM 在最终构型中的方差低于弱贪心方法的方差。R-PDM 和均匀采样的方差最初增加，一种可能的解释是，添加插值节点时， L^1-误差不一定会减小。图 9a 展示了带误差棒的逐点误差，图 9b 展示了逐点方差。均匀采样和弱贪心方法的逐点误差相近，不过贪心方法的方差略低。R-PDM 的逐点误差峰值低于另外两种方法。我们还注意到，逐点方差的上确界较低。定理 5.9 说明了式 (5.3) 中定义的全局误差函数 $\mathscr{G}$ 的收敛性。在 R-PDM 和均匀采样的例子中，数值结果观察到了这种收敛性。我们甚至观察到贪心采样也能收敛，尽管该方法不遵循相同的理论收敛结果。此外，数值例子还为未来的研究指出了方向，因为收敛速率并未说明这些方法的方差。



（A）离散化点处的误差及误差条 fig: 正弦-平均误差　　（B）离散化点处的方差 fig: 正弦-方差

图 9. Example3 的逐点误差和方差

附录 A. Ω及映射 $\oplus$、 N 和 $\mathfrak{S}$的性质

引理 *2.3* 的证明。 通过证明 $(\Omega, \mathfrak{d})$ 是完备且全有界的，可以得出其紧致性。 app: comp性

完备性。 设 $(\eta^n)_{n\in\mathbb{N}}$ 是 Ω 中的一个柯西序列。我们需要证明 $(\eta^n)_{n\in\mathbb{N}}$ 收敛，并首先选取一个候选极限。令 $0 < \varepsilon < \operatorname{diam}(P)$，并令 $i \in \mathbb{N}$。由于 $(\eta^n)_{n\in\mathbb{N}}$ 是一个柯西序列，存在一个 $N \in \mathbb{N}$，使得对所有 $n,\ m \geq N$ 均有

$$\frac{1}{2^i}\bar{\mathfrak{d}}(\eta_i^n, \eta_i^m) \leq \mathfrak{d}(\eta^n, \eta^m) < \frac{\varepsilon}{2^i},$$

也就是说，$(\eta_i^n)_{n\in\mathbb{N}}$ 也是柯西序列。现在我们断言存在一个 $M \in \mathbb{N}$，使得

(*)claim M主张 M 主张 M

$$\begin{cases} \eta_i^n = \phi & \text{for all } n \geq M, \text{ or} \\ \eta_i^n \in P & \text{for all } n \geq M. \end{cases}$$

事实上，对于 $0 < \varepsilon < \operatorname{diam}(\mathrm{P})$，存在一个 $M \in \mathbb{N}$，使得对所有 $n,\ m \geq M$

$$\bar{\mathfrak{d}}(\eta_i^n, \eta_i^m) < \varepsilon.$$

因此，由于对于任意 $\eta_i^n \in P$ 和 $\eta_i^m = \phi$，$\mathfrak{d}(\eta_i^n, \eta_i^m) = \operatorname{diam}(\mathrm{P}) > \varepsilon$，故必有 (*)。由于 $\eta_i^n = \phi$，对于所有 $n \geq M$，$\lim_{n\to\infty} \eta_i^n = \phi$ 蕴含 $\eta_i^n \in P$；对于所有 $n \geq M$，$(\eta_i^n)_{n\geq M}$ 是 $P \subset \mathbb{R}^d$ 中的柯西序列，而 $\mathbb{R}^d$ 是完备的，因此我们得出结论：$(\eta_i^n)_{n\in\mathbb{N}}$ 有极限。

我们将 $(\eta^n)_{n\in\mathbb{N}}$ 的候选极限定义为 $\eta^* = (\eta_1^*, \eta_2^*, \ldots)$，其中对于 $i \in \mathbb{N}$，$\eta_i^* \coloneqq \lim_{n\to\infty} \eta_i^n$。于是 $\eta^* \in \Omega$。接下来我们证明 $\lim_{n\to\infty} \eta^n = \eta^*$。对于所有 $k \in \mathbb{N}$，以及对于所有 $n,\ m \geq N$，

$$\sum_{i=1}^{k} \frac{1}{2^i} \bar{\mathfrak{d}}(\eta_i^n, \eta_i^m) < \varepsilon.$$

令 m 趋于无穷，我们得到

$$\sum_{i=1}^{k} \frac{1}{2^i} \bar{\mathfrak{d}}(\eta_i^n, \eta_i^*) < \varepsilon.$$

接着由 令 k 趋于无穷，我们得到 $\mathfrak{d}(\eta^n, \eta^*) < \varepsilon$，从而完成完备性证明。

全有界性。接下来，我们证明 $(\Omega, \mathfrak{d})$ 是全有界的，即对于每个 $\varepsilon > 0$，都存在一个有限的点集 Ω^ε，使得对于每个 $\eta \in \Omega$，都存在一个 $\gamma \in \Omega^\varepsilon$，满足 $\mathfrak{d}(\eta, \gamma) < \varepsilon$。

设 $\varepsilon > 0$，则存在一个 $K \in \mathbb{N}$，使得 $\sum_{i=K+1}^{\infty} 2^{-i}\mathrm{diam(P)} < \varepsilon/2$。由于 $P \subset \mathbb{R}^d$ 是紧的，存在一个有限集 P^ε，使得对于每个 $p \in P$，都存在一个 $x \in P^\varepsilon$，使得

$$|x - p| < \frac{\varepsilon}{2K}.$$

那么，令 Ω^ε 为

$$\Omega^\varepsilon \coloneqq \left\{\gamma = (\gamma_1, \ldots, \gamma_K, \phi, \ldots) \in \Omega \ :\ \gamma_i \in P^\varepsilon \cup \{\phi\}\right\}.$$

显然，Ω^ε 是有限集。现在，令 $\eta \in \Omega$。若对 $i \in \{1, \ldots, K\}$ 有 $\eta_i = \phi$，则设 $\gamma_i = \phi$；否则设 $\gamma_i = x$，其中 $x \in P^\varepsilon$ 满足 $|\eta_i - x| < \varepsilon/(2K)$。此外，令 $i > K$ 对应的 $\gamma_i = \phi$。则 $\gamma \in \Omega^\varepsilon$，并且

$$\mathfrak{d}(\eta, \gamma) = \sum_{i=1}^{K} \frac{1}{2^i} \bar{\mathfrak{d}}(\eta_i, \gamma_i) + \sum_{i=K+1}^{\infty} \frac{1}{2^i} \bar{\mathfrak{d}}(\eta_i, \gamma_i) < K\frac{\varepsilon}{2K} + \frac{\varepsilon}{2} = \varepsilon.$$

因此，我们得出 $(\Omega, \mathfrak{d})$ 是全有界的，因而是紧的。为证明 N 的连续性，令 $\varepsilon > 0$ 和 $\eta \in \Omega_m$。令 $0 < \delta < \min\{\varepsilon, \mathrm{diam}(P)/2^{m+1}\}$。那么，对所有满足 $\mathfrak{d}(\eta, \gamma) < \delta$ 的 $\gamma \in \Omega$，都有

$$\mathsf{N}(\gamma) - \mathsf{N}(\eta) = 0 < \varepsilon.$$

因此，N 是连续的，因而是可测的。

此外，为证明 $\oplus$ 的连续性，令

$$\mathfrak{d}_{\Omega\times P}((\eta, y), (\gamma, z)) \coloneqq \mathfrak{d}(\eta, \gamma) + |y - z|_2.$$

任取 $(\eta, y) \in \Omega \times P$。令 $\varepsilon > 0$，则对于 $0 < \delta < \min\left(\varepsilon, \mathrm{diam(P)}/2^{\mathsf{N}(\eta)+1}\right)$，我们有：对所有 $(\gamma, z) \in B_\delta((\eta, y))$，均有 $\mathsf{N}(\eta) = \mathsf{N}(\gamma)$，因为 $\delta < \mathrm{diam(P)}/2^{\mathsf{N}(\eta)+1}$。因此，

$$\mathfrak{d}(\eta \oplus y, \gamma \oplus z) = \sum_{i=1}^{\mathsf{N}(\eta)} \frac{1}{2^i} \bar{\mathfrak{d}}(\eta_i, \gamma_i) + \frac{1}{2^{\mathsf{N}(\eta)+1}} |y - z|_2 \leq \sum_{i=1}^{\mathsf{N}(\eta)} \frac{1}{2^i} \bar{\mathfrak{d}}(\eta_i, \gamma_i) + |y - z|_2 < \varepsilon. \qquad \square$$

5.2 的证明。任取 $\varepsilon > 0$，并令 $\eta \in \Omega$。假设 $\mathfrak{d}(\mathfrak{S}(\gamma), \mathfrak{S}(\eta)) < \varepsilon$。我们证明存在某个 $\delta = \delta(\varepsilon) > 0$，使得

$$\mathfrak{d}(\gamma, \eta) < \delta \implies \mathfrak{d}(\mathfrak{S}(\gamma), \mathfrak{S}(\eta)) < \varepsilon.$$

我们首先回顾引理 2.3 的证明：$0 < \delta < \operatorname{diam}(P)/2^{\mathsf{N}(\eta)+1}$，我们有$\mathsf{N}(\gamma) = \mathsf{N}(\eta)$。此外，目前我们假设

性质: 互异性质。 性质: (∗) 不同的

$$\eta_i \neq \eta_j \quad \text{for } i \neq j,\ i, j = 1, \dots, \mathsf{N}(\eta).$$

对于满足性质（∗）的 η，我们定义 $d_{\min} := \min\{|\eta_i - \eta_j| \,|\, i \neq j,\ i, j \in \{1, \dots, \mathsf{N}(\eta)\}\} > 0$。现在令 σ 为满足以下条件的唯一排列

$$\mathfrak{S}(\eta) = (\eta_{\sigma(1)}, \dots, \eta_{\sigma(\mathsf{N}(\eta))}, \phi, \dots).$$

我们声称，当 $\delta > 0$ 充分小时，也有

$$\mathfrak{S}(\gamma) = (\gamma_{\sigma(1)}, \dots, \gamma_{\sigma(\mathsf{N}(\eta))}, \phi, \dots).$$

事实上，取 $\delta < \min\{\varepsilon, d_{\min}\}/2^{\mathsf{N}(\eta)+1}$，可得

$$|\gamma_i - \eta_i| \leq 2^{\mathsf{N}(\eta)-i}|\gamma_i - \eta_i| \leq 2^{\mathsf{N}(\eta)}\mathfrak{d}(\gamma, \eta) < \min\{\varepsilon, d_{\min}\}/2 \qquad \text{for all } i = 1, \dots, \mathsf{N}(\eta),$$

因此可知这些 γ_i' 也互不相同，并且

$$\gamma_{\sigma(i+1)} - \gamma_{\sigma(i)} \geq d_{\min} - |\gamma_{\sigma(i+1)} - \eta_{\sigma(i+1)}| - |\eta_{\sigma(i)} - \gamma_{\sigma(i)}| > 0 \qquad \text{for all } i = 1, \dots, \mathsf{N}(\eta),$$

这就验证了该断言。因此，我们得到

$$\mathfrak{d}(\mathfrak{S}(\gamma), \mathfrak{S}(\eta)) = \sum_{i=1}^{\mathsf{N}(\eta)} \frac{1}{2^i}|\gamma_{\sigma(i)} - \eta_{\sigma(i)}| < \frac{1}{2}\min\{\varepsilon, d_{\min}\} \leq \varepsilon,$$

这就证明了性质（∗）下的断言。

对于一般情形，我们选择满足性质（∗）的 $\tilde{\eta} \in \Omega$，使得对于某个 $\delta_1 < 0$ 有 $\mathfrak{d}(\tilde{\eta}, \eta) < \delta_1 < \varepsilon/2$，并且存在一个排列 σ，使得对所有 $i = 1$、…、$\mathsf{N}(\eta)$，均有 $\eta_{\sigma(i)} = \tilde{\eta}_{\sigma(i)}$，…，$\mathsf{N}(\eta)$。这样，我们像之前那样选择 $\delta_2 > 0$，以便对 $\mathfrak{d}(\gamma, \tilde{\eta})$ 得到

$$\mathfrak{d}(\mathfrak{S}(\gamma), \mathfrak{S}(\eta)) \leq \mathfrak{d}(\mathfrak{S}(\gamma), \mathfrak{S}(\tilde{\eta})) + \mathfrak{d}(\mathfrak{S}(\tilde{\eta}), \mathfrak{S}(\eta)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

从而完成证明。 □

Appendix B. Well-posedness

我们证明，式（2.1）中给出的、以 L 为生成元的过程存在。在此之前，我们注意到 L 可以表示为 app: 定义良好性

$$LF(\eta) = \int_\Omega \big[F(\sigma) - F(\eta)\big]\, \kappa(\eta, d\sigma),$$

其中，对每个 $(\eta, A) \in \Omega \times \mathcal{B}_\Omega$,

$$\kappa(\eta, A) := \int_P \delta_{\eta \oplus y}(A)\, \lambda(\eta, dy) = \int_P \mathbf{1}_A(\eta \oplus y)\, \lambda(\eta, dy).$$

接下来要证明，κ 是一个定义良好且有界的转移核。

引理 B。 **1.** 上述定义的映射 $\kappa : \Omega \times \mathcal{B}_\Omega \to [0, +\infty)$ 是一个有界的转移核 *l.*

证明。 κ 的有界性直接由 λ的有界性得到。事实上，对于每个 $\eta \in \Omega$，我们有 $\kappa(\eta, \Omega) = \lambda(\eta, P) = 1$。

为证明 κ 是一个转移核，我们首先注意到，对于任意 $A \in \mathcal{B}_\Omega$，简单函数 $\Omega \times P \ni (\eta, y) \mapsto \mathbf{1}_A(\eta \oplus y)$ 是博雷尔可测的，因为它是两个博雷尔可测映射的复合。根据标准的单调类论证，我们随后得到 $\Omega \ni \eta \mapsto \kappa(\eta, A)$是博雷尔可测的。

为证明对于每个 $\eta \in \Omega$都有 $\kappa(\eta, \cdot) \in \mathcal{P}(\Omega)$，我们证明测度的一个等价定义。显然，对于不相交的集合 A, $B \in \mathcal{B}_\Omega$, 有 $\kappa(\eta, \emptyset) = 0$和 $\kappa(\eta, A \cup B) = \kappa(\eta, A) + \kappa(\eta, B)$。现在令$\{A_i\}_{n \in \mathbb{N}} \subset \mathcal{B}_\Omega$ 为任意满足 $\cup_{i \in \mathbb{N}} A_i \in \mathcal{B}_\Omega$的可测集合递增族。根据单调收敛定理和狄拉克测度的下连续性，我们得到

$$\begin{aligned}\lim_{i\to\infty} \kappa(\eta, A_i) &= \lim_{i\to\infty} \int_P \mathbf{1}_{A_i}(\eta \oplus y)\, \lambda(\eta, dy) = \int_P \lim_{i\to\infty} \mathbf{1}_{A_i}(\eta \oplus y)\, \lambda(\eta, dy) \\ &= \int_P \lim_{i\to\infty} \delta_{\eta\oplus y}(A_i)\, \lambda(\eta, dy) = \int_P \delta_{\eta\oplus y}(\cup_{i\in\mathbb{N}} A_i)\, \lambda(\eta, dy) = \kappa(\eta, \cup_{i\in\mathbb{N}} A_i).\end{aligned}$$

结合 $\kappa(\eta, \Omega) = 1$, 这意味着对于每个 $\eta \in \Omega$, $\kappa(\eta, \cdot) \in \mathcal{P}(\Omega)$。 □

致谢

本研究是一个项目的一部分，该项目获得了欧盟“地平线2020”研究与创新计划下欧洲研究理事会（ERC）的资助（资助协议编号：818473）。作者感谢 Karen Veroy-Grepl 提供的宝贵意见和讨论。作者承认使用了 Grammarly 和 ChatGPT 对文本进行润色，以改进拼写、语法和整体文风。

参考文献

[1] M. Barrault 等。“一种‘经验插值’方法：应用于偏微分方程的高效降阶基离散化”。载于：*ComptesRendusMathematique*339.9（2004），第667–672页。[2] A. R. Barron 等。“通过贪心算法进行近似与学习”。载于：*The Annals ofStatistics* 36.1（2008），第64–94页。[3] P. Binev 等。“降阶基方法中贪心算法的收敛速率”。载于：*SIAM Journal on MathematicalAnalysis*43.3（2011），第1457–1472页。[4] R. Campagna 等。“通过指数–多项式样条进行学习的贪心算法”。载于：*arXivpreprintarXiv:2109.14299* （2021）。[5] P. Chen、A. Quarteroni 和 G. Rozza。“加权经验插值方法：先验收敛分析与应用”。载于：*ESAIM*：数学建模与数值分析 48.4（2014），第943–953页。[6] Y. Chen 等。“调和 Maxwell 方程的经认证降阶基方法与输出界”。载于：*SIAM* 科学计算杂志 32.2（2010），第970–996页。[7] A. Cohen 等。“使用随机训练集进行降阶基贪心选择”。载于：*ESAIM*：数学建模与数值分析54.5（2020），第1509–1524页。

[8] S. Deparis。“通过自然范数方法计算参数依赖的纳维–斯托克斯方程的降阶基误差界”。载于：*SIAM* 数值分析杂志46.4（2008），第2039–2067页。[9] R. A. DeVore 和 V. N. Temlyakov。“关于贪心算法的一些评述”。载于：计算数学进展 5.1（1996），第173–187页。[10] S. N. Ethier 和 T. G. Kurtz。马尔可夫过程：刻画与收敛。约翰·威立父子公司，2009年。[11] B. Haasdonk、M. Dihlmann 和 M. Ohlberger。“一种基于参数空间自适应网格、用于参数化模型降阶的训练集与多基生成方法”。载于：数学与计算机动力系统建模17.4（2011），第423–442页。[12] J.S. Hesthaven、B. Stamm 和 S. Zhang。“适用于高维参数空间的高效贪心算法及其在经验插值和降阶基方法中的应用*”。载于：*ESAIM*：数学建模与数值分析 48.1（2014），第259–283页。[13] D. B. P. Huynh 和 A. T. Patera。“应力强度因子的降阶基近似与后验误差估计”。载于：国际工程数值方法杂志 72.10（2007），第1219–1259页。[14] J. Jiang、Y. Chen 和 A. Narayan。“通过自适应构造代理训练集实现的离线增强降阶基方法”。载于：科学计算杂志 73（2017），第853–875页。[15] R. B. Kelman 和 T. J. Rivlin。“反常积分的被积函数有界或趋于零的条件”。载于：美国数学月刊 67.10（1960），第1019–1022页。[16] Y. Li。“经验插值方法与切比雪夫贪心算法的新分析”。载于：*SIAM* 数值分析杂志 63.2（2025），第931–948页。[17] Y. Maday 等。“一种通用的多用途插值过程：魔点”。载于：（2007）。[18] B. Mirzasoleiman 等。“比懒惰贪心更懒惰”。载于：人工智能促进协会会议。第29卷，第1期，2015年。[19] F. Negri 等。“参数化椭圆最优控制问题的降阶基方法”。载于：*SIAM* 科学计算杂志 35.5（2013），A2316–A2340。[20] N. C. Nguyen、K. Veroy 和 A. T. Patera。“参数化偏微分方程的经认证实时求解”。载于：材料建模手册（2005），第1523–1558页。[21] E. Nielen、O. Tse 和 K. Veroy。“多胞体划分方法：一种适用于高维参数问题的可扩展采样方法”。载于：*SIAM* 科学计算杂志 47.6（2025），B1424–B1449。[22] G. Rozza。“参数化几何中势流的降阶基近似与误差界”。载于：计算物理通讯 9.1（2011），第1–48页。[23] G. Rozza、D. B. P. Huynh 和 A. T. Patera。“仿射参数化椭圆强制偏微分方程的降阶基近似与后验误差估计：在输运和连续介质力学中的应用”。载于：*Archives ofComputational Methodsin Engineering* 15.3（2008），第229–275页。

[24] G. Rozza 和 K. Veroy。《参数化区域中 Stokes 方程约化基方法的稳定性》。载于：计算机在应用力学与工程中的方法196.7（2007），第 1244–1260 页。[25] G. Rozza 等。《传热问题的降阶基方法与后验误差估计器》。载于：传热夏季会议。HT2009-88211。2009 年，第 753–762 页。[26] R. Schaback。《基于节点数据的泛函贪心稀疏线性近似》。载于：数值算法 67（2014），第 531–547 页。[27] S. Sen。《多参数热传导问题的约化基近似与后验误差估计》。载于：数值传热，*B* 辑：基础 54.5（2008），第 369–389 页。[28] D. Ucinski。分布参数系统辨识的最优测量方法。CRC出版社，2004 年。[29] K. Urban、S. Volkwein 和 O. Zeeb。《使用非线性优化进行贪心采样》。载于：建模与计算约简的降阶方法 （2014），第 137–157 页。[30] K. Veroy 等。《参数化非强制与非线性椭圆偏微分方程约化基近似的后验误差界》。载于：第 *16* 届 *AIAA* 计算流体动力学会议。2003 年，第 3847 页。[31] T. Wenzel、F. Marchetti 和 E. Perracchione。《用于优化贪心方案的数据驱动核设计：机器学习视角》。载于：*SIAM* 科学计算期刊 46.1（2024），第 C101–C126 页。[32] D. Wu、C-T. Lin 和 J. Huang。《使用贪心采样进行回归的主动学习》。载于：信息科学 474（2019），第 90–105 页。