

散乱数据拟合的自适应分片逆尺度空间算法

钟轶君¹, 李崇君²

(1. 浙江理工大学理学院数学科学系, 浙江 杭州 310018;
2. 大连理工大学数学与科学学院, 辽宁 大连 116024)

摘 要: 散乱数据拟合(逼近)是在信号处理、计算机图形学等领域中被广泛研究的问题, 近些年, 利用优化方法获得散乱数据的稀疏表示逼近解也成为了优化和曲面重构交叉领域的热点。基于由 B 样条生成的 PSI 空间中的散乱点曲面拟合问题和分片稀疏的联系, 将分片稀疏性引入到 Bregman 逆尺度空间算法(ISS)中, 提出一种自适应的分片逆尺度空间(aP_ISS)算法, 处理散乱数据的曲面拟合问题。通过对逆尺度空间系统分片符号一致性分析, 得到了自适应分片逆尺度空间系统的性能保证定理和避免了 aP_ISS 参数的选取。应用到散乱点曲面重构问题上的数值实验结果表明, 该算法不仅可以有效拟合曲面, 还能够较好保护分片稀疏性。

关 键 词: 散乱数据拟合; 分片稀疏; 逆尺度空间; 稀疏优化

中图分类号: TP 391

DOI: 10.11996/JGJ.2095-302X.2020010018

文献标识码: A

文章编号: 2095-302X(2020)01-0018-09

Adaptive piecewise inverse scale space algorithm for scattered data fitting

ZHONG Yi-jun¹, LI Chong-jun²

(1. Department of Mathematical Sciences, Zhejiang Sci-Tech University, Hangzhou Zhejiang 310018, China;

2. School of Mathematical Sciences, Dalian University of Technology, Dalian Liaoning 116024, China)

Abstract: Scattered data reconstruction has been widely studied in the fields of signal processing and computer graphics. Moreover, in recent years, to obtain sparse representation approximations of scattered data by means of sparse optimization method has also become a hot spot in the cross-domain of optimization and surface reconstruction. In this paper, we establish the connection between surface fitting of scattered data and the piecewise sparseness in PSI space generated by a B-spline, and introduce the piecewise sparsity to the Bregman inverse scale space (ISS) algorithm. In addition, an adaptive piecewise ISS algorithm is established to solve the scattered data reconstruction problem. Through the analysis of the piecewise symbolic consistency, the performance guarantee of adaptive piecewise ISS system is obtained in this paper and the selection of aP_ISS parameters can be avoided. Numerical experimental results applied to the surface reconstruction of scattered data show that, this algorithm can not only effectively fit the surface, but also protect the piecewise sparsity of coefficient of the surface.

Keywords: scattered data reconstruction; piecewise sparsity; inverse scale space; sparse optimization

收稿日期: 2019-07-21; 定稿日期: 2019-09-30

基金项目: 国家自然科学基金项目(11901529, 11871137, 11572081); 浙江理工大学科研启动基金项目(19062118-Y)

第一作者: 钟轶君(1988-), 女, 江西九江人, 博士, 讲师。主要研究方向为曲面稀疏表示、稀疏优化、散乱数据造型与样条。

E-mail: zhongyijun@zstu.edu.cn

散乱数据拟合(函数拟合/逼近)是一个在信号处理、计算机图形学中被广泛研究的问题。其通过给定的散乱的、带噪音的点集来拟合一个较好的曲线或曲面, 需要解决 2 个基本问题: ①如何选择一个简单的函数空间作为描述对象函数的逼近空间; ②其空间要有很好的逼近度。为此, 随着散乱点曲面拟合工具多样性的不断探索, 有学者开始利用探索多层结构性质来改进逼近方法, 如 LEE 等^[1]提出的基于 B 样条的多层逼近方法; CASTANO 等^[2]提出了基于小波的多层正则化方法等。JOHNSON 等^[3]进一步探讨了在平移不变空间(principal shift invariant, PSI)中基于 B 样条和小波的散乱点曲面拟合问题。以上方法均力求在所构建的函数类中与研究对象函数本身误差最小。对于第②个问题, 学者们希望在其空间中得到散乱数据的稀疏表示解, 因此推动了稀疏逼近和优化方法在函数拟合领域的广泛应用。随着深层挖掘曲面中的稀疏性, 几何建模中很多的启发式方法得到了进一步发展和推广。近些年, 信号的稀疏性研究也取得了丰硕的成果, 其中细化研究信号的稀疏分布(非零元素的分布)成为了热点, 如 HUANG^[4]提出了结构化稀疏的概念, 即一类非零元具有一定结构性的稀疏向量, 如广泛研究的非零元是成块出现的块稀疏^[5-9], 在统计中也称为 Group Lasso^[10-17], 除此之外还有树稀疏^[18]、图稀疏^[19]。刘翼鹏^[19]讨论了向量服从全局稀疏、局部分布不同的凸优化模型, 讨论了全局稀疏局部稠密, 即非零元素呈簇状分布的稀疏向量。提取出了可以描述整个信号的结构性信息, 有利于使用更少的样本和运算量来恢复信号, 且能够决定信号是否可以有效地得到恢复。对于不同的实际场景, 信号的稀疏表示向量具有不同的结构, 更好地研究这些结构有利于理解和解决实际问题。随着对稀疏性的细化探索, 学者们尝试在曲面重构, 尤其是在 PSI 空间中的曲面重构问题所具有的结构型稀疏表示进行研究。HAO 等^[20]提出了利用凸优化模型来考虑 PSI 空间中基于 B 样条的散乱点曲面拟合问题, 利用多块 LASSO (MLASSO)得到散乱点曲面拟合问题的稀疏解; 文献[21]得到了曲面在 PSI 空间中每一层 B 样条上的稀疏表示解。基于 MLASSO 的散乱点曲面拟合问题不仅可以得到稀疏解(稀疏表示系数)而且拟合效果较好, 但其多个正则化参数的选取是一个待解决的问题。OSHER 等^[22]提出 Bregman 逆尺度空间算法(inverse scale space, ISS), 即

$$\partial_t p_t = A^T(b - Ax(t)), p_t \in \partial \|x(t)\|_1 \quad (1)$$

此方法是贪婪算法和凸优化算法的一种结合, 不仅规避了凸优化算法中参数的选取问题而且运行速率较快。之后 LI 和 ZHONG^[23]提出了一种带有删除机制的分片逆尺度空间算法, 核心是将分片稀疏性引入到 ISS 算法中。在稀疏恢复问题上, 相较于 ISS 算法, 该算法能够更好地保护信号的小尺度非零元素, 在应用于散乱数据拟合问题中得到了较好的拟合效果。分片稀疏性是一类描述向量中非零元素散乱分布现象的稀疏性。DONOHO 和 HUO^[24]针对片稀疏信号对应观测矩阵具有的分块结构, 引入分块矩阵互相干(mutual coherence)条件, 可将分块正交矩阵的精确恢复信号的稀疏度理论上界推广到一般分块矩阵之中, 改进了稀疏信号可精确恢复的充分性条件。其结果不仅给出了分片稀疏恢复的理论保证, 也提升了 L0 模型和压缩感知恢复整体稀疏信号的可信度。向量的分片稀疏结构所对应的观测矩阵 A 的分块结构恰好符合 PSI 空间基于 B 样条所生成基函数的层数, 因此引入分片稀疏性到稀疏优化算法(模型)中, 并分析散乱点曲面拟合问题是有意义的。

然而分片 ISS 算法(P_ISS)^[23]也有其缺陷, 算法中的删除机制增加了新的参数选取问题。目前考虑基于散乱点的曲面拟合的稀疏表示的文献较少。基于此, 本文提出一种自适应的分片逆尺度空间算法(adaptive piecewise ISS, aP_ISS), 由此得到散乱点曲面拟合问题的稀疏表示。此算法不仅规避了带有删除机制的分片逆尺度算法中需要人工选取参数的问题, 而且通过对“分片符号一致性”的分析, 使得算法具有自适应性, 可更好地应用到散乱点曲面逼近问题中。

本文主要贡献如下:

(1) 通过将分片稀疏的理念引入到散乱数据拟合问题中, 拓展了稀疏技术在函数逼近中的应用。

(2) 改进了文献[22]中的分片逆尺度空间算法, 规避了算法中的删除机制。利用“分片符号一致性”使得算法更具自适应性。

(3) 稀疏算法作为高效算法已经成功地应用到了曲面拟合问题中, 本文建立了一个稀疏表示系数的先验信息, 从而可以更好拟合曲面的细节信息, 达到更好地逼近效果。与已有的稀疏重建算法相比, 不仅有更松弛的理论保证并且拟合效果更好。

1 预备知识

1.1 散乱数据曲面拟合问题

典型的散乱点曲面拟合问题是指：存在一个散乱点集合 $X = \{x_1, x_2, \dots, x_n\} \subset \Omega \subset \mathbf{R}^d$ 和相对应的函数值 $f|_X = \{f_1, f_2, \dots, f_n\}$ ，且需要找到一个属于空间 H 的函数 g 能够拟合给定的数据 $\{(x_k, f_k)\}_{k=1}^n$ 。经典的光滑样条散乱点重构模型为

$$\min \sum_{k=1}^n (g(x_k) - f_k)^2 + \alpha \|g\|_{H^m}^2 \quad (2)$$

由于 g 属于 Beppo-Levi 空间，所以当数据点规模变大时计算量也会变大。为了解决该问题，文献[3]考虑在一个由紧致支撑函数移位扩张生成的空间(称为 PSI 空间)中求解以上正则化问题。利用 PSI 空间不仅避免了计算量增大的问题，也由于其和小波、B 样条的联系使得稀疏逼近数据函数成为可能。

记 $\mathcal{H} = \bigcup_{i=1}^N H_i$ 为逼近空间，其中 $H_i (\subseteq H_{i+1})$ 为 B 样条函数 ϕ_i 生成的 PSI 空间，此时待重构的曲面 g 为

$$g = \sum_{i=1}^N g_i, \quad g_i \in H_i, \quad g_i = \sum_{j=1}^m c_j^i \phi_j^i \quad (3)$$

其中， c_j^i 为系数； $H_i = \text{span}\{\phi_j^i, j=1, \dots, m\}$ 。设 h 是 PSI 空间的尺度参数，指标集 K 为

$$K = \left\{ k \in \mathbf{Z}^2 : \text{supp} \left(\phi \left(\frac{2^s \cdot}{h} - k \right) \right) \cap \Omega \neq \emptyset \right\} \quad (4)$$

利用分块矩阵 $A = [A_1, \dots, A_N]$ 表示由 ϕ_j^i 定义的观测矩阵

$$A_i(s, k) = \phi \left(\frac{2^i x_s}{h} - k \right), \quad k \in K \quad (5)$$

向量 $\mathbf{b}^\sigma$ 表示被噪音污染的数据 $\{f_i\}$ ，则曲面拟合问题 $g(x_i) \approx f_i$ 可以记为 $\mathbf{b}^\sigma = A\mathbf{x} + \mathbf{e}$ ，其中 $\mathbf{x}$ 为曲面表示系数； $\mathbf{e}$ 为噪音。

1.2 符号说明

定义 1^[24]。观测矩阵 A 的列互相干为

$$\mu := \max_{i,j} |\langle \mathbf{a}_i, \mathbf{a}_j \rangle| \quad (6)$$

其中， $\mathbf{a}_i$ 为矩阵 A 的第 i 列。

定义 2^[23]。设矩阵 $A = [A_1, \dots, A_N]$ 是 N 个矩阵的拼接，定义第 i 块的子阵互相干为

$$\mu^{i,i} = \max_i \max_{k \neq l} |\langle \mathbf{a}_k^i, \mathbf{a}_l^i \rangle| \quad (7)$$

其中， $\mathbf{a}_k^i, \mathbf{a}_l^i \in A_i, i=1, \dots, N$ 。

命题 1^[23]。第 i 块的子阵互相干 $\mu^{i,i}$ 满足

$$0 \leq \mu^{i,i} = \alpha_i \mu \leq \mu \quad (8)$$

其中， $\alpha_i \in [0, 1]$ 为衡量第 i 个子矩阵块内互相干与矩阵整体互相干比例关系的参数。

Oracle 估计^[21]是最小二乘解在真实支撑集 S 上的部分，oracle 的非零元素可表达为

$$\hat{\mathbf{x}}_S = (A_S^T A_S)^{-1} A_S^T \mathbf{b}^\sigma = \mathbf{x}_S^* + (A_S^T A_S)^{-1} A_S^T \mathbf{e} \quad (9)$$

分片 oracle 估计^[22]是 oracle 估计限制在子空间 S_i 上的部分

$$\hat{\mathbf{x}}_{S_i} = \mathbf{x}_{S_i}^* + (A_{S_i}^T A_{S_i})^{-1} A_{S_i}^T \mathbf{e} + (A_{S_i}^T A_{S_i})^{-1} A_{S_i}^T \sum_{j \neq i} A_{S_j} (\mathbf{x}_{S_j}^* - \hat{\mathbf{x}}_{S_j}) \quad (10)$$

其中， $S = (S_i)_{i=1}^N$ 。

1.3 逆尺度空间算法

文献[25]中提出了一个自适应的逆尺度空间算法(adaptive inverse scale space, aISS)，实现算法步骤如下：

输入： A ， $\mathbf{b}^\sigma$ ，threshold ≥ 0 。

输出： $\mathbf{x}(t_k)$ 。

步骤 1. 初始化 $t_1 = 1/\|A^T \mathbf{b}^\sigma\|_\infty$ ， $\mathbf{p}(t_1) = t_1 A^T \mathbf{b}^\sigma$ ， $I_1 = \{i | p_i(t_1) = 1\}$ 。

步骤 2. 计算 $\mathbf{x}(t_k)$ 为问题

$$\begin{cases} \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}^\sigma\|_2^2 \\ \text{s.t. } (p_{t_{k+1}})_i x_i \geq 0, \forall i \in S_{k+1} \\ x_j = 0, \forall j \in T_{k+1} \end{cases} \quad (11)$$

在 $S_k = I_k$ 上的解。

步骤 3. 更新时间变量为

$$\begin{aligned} t_{k+1} &= \min \{t > t_k \mid \exists j, |p_j(t)| = 1 \\ &\quad x_j(t_k) = 0, p_j(t) \neq p_j(t_k)\} \end{aligned} \quad (12)$$

步骤 4. 更新次梯度为

$$\mathbf{p}_j(t) = \mathbf{p}_j(t_k) + (t - t_k) \mathbf{e}_j \cdot A^T (\mathbf{b}^\sigma - A\mathbf{x}(t_k)) \quad (13)$$

其中， $t = t_{k+1}$ ， $\mathbf{e}_j \in \mathbf{R}^n$ 为第 j 个分量为 1 其余分量为 0 的向量。

步骤 5. 更新指标集为

$$I_{k+1} = \{i \mid p_i(t_{k+1}) = 1\} \quad (14)$$

文献[21]中给出了如下 ISS 系统(1)的理论保证条件。

假设 1^[21]。限制强凸性 (restricted strong convexity)：存在 $\gamma \in (0, 1]$ ，使得 $A_S^T A_S \geq \gamma I$ 。

假设 2^[21]。不可表示条件或精确恢复条件 (irre-presentable condition or exact recovery condition)：

存在 $\eta \in (0, 1)$, 使得 $\|A_T^T A_S^\dagger\|_\infty \leq 1 - \eta$, 其中 $A_S^\dagger = A_S(A_S^T A_S)^{-1}$.

定理 1^[21]. 设假设 1 和 2 均成立, 那么存在一个时刻 τ 使得 Bregman ISS 路径(1)在时刻 τ 之前没有错误选择, 即 $\forall t \leq \tau, \text{supp}(\mathbf{x}(t)) \subseteq \text{supp}(\mathbf{x})$ 以大于

$$1 - \frac{2}{n\sqrt{\pi \log_{10} n}} \text{ 的概率成立. 若 } x_{\min} \geq \max \left\{ \frac{4\sigma}{\gamma^{1/2}}, \frac{8\sigma(2 + \log_{10} s) \left(\max_{j \in T} \|A_j\|_2 \right)}{\gamma\eta} \right\} \sqrt{\log_{10} n} \quad (15)$$

成立, 则以超过 $1 - \frac{2}{n\sqrt{\pi \log_{10} n}}$ 的概率有

$\text{sign}(\mathbf{x}(\tau)) = \text{sign}(\mathbf{x})$, 即 Bregman ISS 求解的最优解达到了符号一致性.

1.4 分片稀疏

对于给定的向量

$$\mathbf{x} = (\underbrace{x_1, \dots, x_{d_1}}_{\mathbf{x}_1^T}, \underbrace{x_{d_1+1}, \dots, x_{d_1+d_2}}_{\mathbf{x}_2^T}, \dots, \underbrace{x_{n-d_N+1}, \dots, x_n}_{\mathbf{x}_N^T})^T \quad (16)$$

其中, $n = \sum_{i=1}^N d_i$, $s_i = \|\mathbf{x}_i\|_0$ ($i=1, \dots, N$). 对于向量而言有 3 种常用的稀疏性:

(1) 全局稀疏. $\mathbf{x}$ 称为 s -稀疏的若 $\mathbf{x}$ 至多含有

$$\|\mathbf{x}\|_0 = \sum_{i=1}^N \|\mathbf{x}_i\|_0 \leq s \text{ 个非零元素;}$$

(2) 块稀疏^[4-9]. $\mathbf{x}$ 称为块 s -稀疏的若 $\mathbf{x}$ 至多含有 s 个含有非零元素的块, 即块 l^0 范数 $\|\mathbf{x}\|_{2,0} =$

$$\sum_{i=1}^N I(\|\mathbf{x}_i\|_2) \text{ 不超过 } s;$$

(3) 分片稀疏^[23]. 向量 $\mathbf{x} = (\mathbf{x}_1^T, \dots, \mathbf{x}_N^T)^T \in \mathbf{R}^n$ 可以分成 N 个部分, 且每部分 $\mathbf{x}_i^T \in \mathbf{R}^{n_i}$ 都是一个稀疏向量, 若已知 $s_i = \|\mathbf{x}_i\|_0$ ($i=1, \dots, N$), 则称 $\mathbf{x}$ 是分片 $(s_1, \dots, s_N)$ -稀疏向量.

1.5 分片逆尺度空间算法(P_ISS)

分片逆尺度空间系统(piecewise inverse scale space, P_ISS)^[22] 旨在恢复分片稀疏向量 $\mathbf{x} = (\mathbf{x}_1^T, \dots, \mathbf{x}_N^T)^T \in \mathbf{R}^n$ 中每一个子向量 $\mathbf{x}_i^T \in \mathbf{R}^{n_i}$ 的支集, 即 P_ISS 方法分别且同时求解 $\mathbf{x}_i$, 再由所有的 $\mathbf{x}_i$ 共同构成 $\mathbf{x}$, 即

$$\begin{cases} \partial_{t_i} \mathbf{p}_i(t_i) = A_i^T (\mathbf{b}^\sigma - \sum_{i=1}^N A_i \mathbf{x}_i), & i=1, \dots, N \\ \mathbf{p}_i(t_i) \in \partial(\|\mathbf{x}_i\|_1), \end{cases} \quad (17)$$

算法 1^[22]. 分片逆尺度空间算法

输入: 观测矩阵 $\mathbf{A} = [\mathbf{A}_1, \dots, \mathbf{A}_N]$, 观测向量 $\mathbf{b}^\sigma$, 噪音程度参数 σ .

输出: 分片稀疏向量 $\mathbf{x} = (\mathbf{x}_1^T, \dots, \mathbf{x}_N^T)^T \in \mathbf{R}^n$.

步骤 1. 初始化.

$$\begin{aligned} t_i^1 &= 1/\|A_i^T \mathbf{b}^\sigma\|_\infty \\ \mathbf{p}_i(t_i^1) &= t_i^1 A_i^T \mathbf{b}^\sigma \end{aligned} \quad (18)$$

$$\mathcal{I}_i^1 = \{j : |\mathbf{p}_i(t_i^1)(j)| = 1\}$$

步骤 2. 最小二乘. 更新 $(\mathbf{x}_1^T(t_1^k), \dots, \mathbf{x}_N^T(t_N^k))^T$ 为

$$\begin{cases} \min_{(\mathbf{x}_1^T, \dots, \mathbf{x}_N^T)} \|\mathbf{A}_1 \mathcal{P}_{\mathcal{I}_1} \mathbf{x}_1 + \dots + \mathbf{A}_N \mathcal{P}_{\mathcal{I}_N} \mathbf{x}_N - \mathbf{b}^\sigma\|^2 \\ \text{s.t. } \mathbf{x}_i(t_i^k) \mathbf{p}_i(t_i^k) \geq 0, i=1, \dots, N \end{cases} \quad (19)$$

的解. 其中 $\mathcal{I}$ 为指标集; $\mathcal{P}_{\mathcal{I}}$ 为在指标集 $\mathcal{I}$ 上的投影算子.

步骤 3. 并行时间路径更新.

$$t_i^{k+1} = \min \{t_i : t_i > t_i^k, \exists l : |\mathbf{p}_i(t_i)(l)| = 1\} \quad (20)$$

$$\mathbf{x}_{i(l)}(t_i^k) = 0, \mathbf{p}_{i(l)}(t_i) \neq \mathbf{p}_{i(l)}(t_i^k)\}$$

步骤 4. 并行的次梯度向量更新.

$$\mathbf{p}_{i(j)}(t_i^{k+1}) = \mathbf{p}_{i(j)}(t_i^k) + (t_i^{k+1} - t_i^k) \mathbf{e}_j \cdot A_i^T \left(\mathbf{b}^\sigma - \sum_{i=1}^N A_i \mathbf{x}_i(t_i^k) \right) \quad (21)$$

步骤 5. 并行的指标集更新.

$$\mathcal{I}_i^{k+1} = \{j : |\mathbf{p}_{i(j)}(t_i^{k+1})| = 1\} \quad (22)$$

步骤 6. 删除机制. 设置向量

$$\mathbf{x}^k = (\mathbf{x}_1^T(t_1^k), \dots, \mathbf{x}_N^T(t_N^k))^T \quad (23)$$

满足不等式

$$x_j \leq C \frac{4\sigma}{\sqrt{\gamma}} \sqrt{\log_{10} n}, \quad j=1, \dots, n \quad (24)$$

的分量为 0. 引入分片稀疏性后提出的带删除机制的分片 Bregman 逆尺度空间(P_ISS)算法. 因该算法结构类似于正交匹配算法, 所以其运行速率较快, 又因求解的是分片凸优化问题, 故其收敛到 l^1 稀疏解. 数值实验表明相比于 ISS 算法, 分片 Bregman 逆尺度空间算法能够更好地保护小尺度元素不被噪音混淆.

2 自适应分片逆尺度空间算法(aP_ISS)

2.1 P_ISS 算法中删除机制的缺陷

在改良的贪婪算法中, 比如压缩采样匹配追踪算法 (compressive sampling MP, CoSaMP)^[26], 利用删除机制对算法当前恢复的解进行自适应

的修正是较为常用且有效的方法。CoSaMP 在算法最后一步保留最大的 $s(s=\|\mathbf{x}\|_0)$ 个非零元素,其余的元素均设置为 0 以克服类似 StOMP 算法^[27]在支集添加冗余元素的缺陷。在算法 P_ISS 迭代过程中,由于分片稀疏向量 $\mathbf{x}$ 整体到达符号一致性^[22]的时刻 $\tau = \max_{i=1,\dots,N} \{\tau_1, \dots, \tau_N\}$, 其中 τ_i 为第 i 片到达符号一致性的时刻。即所有的“片”到达了符号一致性算法才会停止更新支集。在所有的 N 个子向量都到达符号一致性之前,早到达符号一致性的“片”会由于不断的迭代产生错误或冗余的非零元素,所以在 P_ISS 算法中设置删除机制可以有效地避免冗余元素的产生。但是,删除机制中引入的参数 C 使得算法增加了选取问题,即参数 C 与数据密切相关,需要经验性选取,消耗大量时间。

2.2 分片符号一致性

本文考虑对算法进行分片符号一致性的改进,即当某一“片”到达符号一致性时,停止此“片”的支集更新。从而保证在最小二乘计算中,该“片”不会随着晚到达符号一致性的“片”的更新而增加冗余的元素。该算法自行地对每一“片”实行分片符号一致性的“监督”,使得算法自适应性更强。

2.3 自适应分片逆尺度空间算法(aP_ISS)

算法 2. 自适应分片逆尺度空间算法

输入: 观测矩阵 $\mathbf{A}=[\mathbf{A}_1, \dots, \mathbf{A}_N]$, 观测向量 $\mathbf{b}^\sigma$,

噪音程度参数 σ 。

输出: $\mathbf{x}=(\mathbf{x}_1^T, \dots, \mathbf{x}_N^T)^T$ 。

步骤 1. 初始化。

$$\begin{aligned} t_i^1 &= 1 / \|\mathbf{A}_i^T \mathbf{b}^\sigma\|_\infty \\ \mathbf{p}_i(t_i^1) &= t_i^1 \mathbf{A}_i^T \mathbf{b}^\sigma \\ \mathcal{I}_i^1 &= \{j : |p_{i(j)}(t_i^1)| = 1\} \end{aligned} \quad (25)$$

步骤 2. 最小二乘。更新 $(\mathbf{x}_1^T(t_1^k), \dots, \mathbf{x}_N^T(t_N^k))^T$ 为

$$\begin{cases} \min_{(\mathbf{x}_1^T, \dots, \mathbf{x}_N^T)} \|\mathbf{A}_1 \mathcal{P}_{\mathcal{I}_1} \mathbf{x}_1 + \dots + \mathbf{A}_N \mathcal{P}_{\mathcal{I}_N} \mathbf{x}_N - \mathbf{b}^\sigma\|^2 \\ \text{s.t. } \mathbf{x}_i(t_i^k) \mathbf{p}_i(t_i^k) \geq 0, i=1, \dots, N \end{cases} \quad (26)$$

的解。其中, $\mathcal{I}$ 为指标集; $\mathcal{P}_{\mathcal{I}}$ 为在指标集 $\mathcal{I}$ 上的投影算子。

步骤 3. 并行时间路径更新。

$$\begin{aligned} t_i^{k+1} &= \min \{t_i : t_i > t_i^k, \exists l : |p_{i(l)}(t_i)| = 1 \\ &\quad \mathbf{x}_{i(l)}(t_i^k) = 0, p_{i(l)}(t_i) \neq p_{i(l)}(t_i^k)\} \end{aligned} \quad (27)$$

步骤 4. 并行的次梯度向量更新。

$$p_{i(j)}(t_i^{k+1}) = p_{i(j)}(t_i^k) + (t_i^{k+1} - t_i^k) e_j \cdot \mathbf{A}_i^T \left(\mathbf{b}^\sigma - \sum_{i=1}^N \mathbf{A}_i \mathbf{x}_i(t_i^k) \right) \quad (28)$$

步骤 5. 并行的自适应的指标集更新。若第 i 片未达到符号一致性, 则

$$\mathcal{I}_i^{k+1} = \{j : |p_{i(j)}(t_i^{k+1})| = 1\} \quad (29)$$

否则 $\mathcal{I}_i^{k+1} = \mathcal{I}_i^k$ 。

实际算法操作中, 通过算法当前迭代向量和 oracle 向量的符号一致性, 即考虑第 i 片的向量 $\mathbf{x}_i$ 和分片 oracle 估计向量 $\hat{\mathbf{x}}_{S_i}$ 是否达到了符号一致性

$$\text{sign}(\mathbf{x}_i) = \text{sign}(\hat{\mathbf{x}}_{S_i}) \quad (30)$$

并衡量是否更新指标集。下面给出 aP_ISS 的性能保证定理。首先给出一些相关的引理。

引理 1. 矩阵 $\mathbf{A}_S = [\mathbf{A}_{S_1}, \dots, \mathbf{A}_{S_N}]$ 的奇异值 Λ 满足

$$\begin{aligned} 1 - (s_l - 1)\alpha_l \mu - (s - s_l)\mu &\leq \Lambda^2 \leq \\ 1 + (s_l - 1)\alpha_l \mu + (s - s_l)\mu \end{aligned} \quad (31)$$

其中,

$$\mathcal{I} = \left\{ i_0 : (s_{i_0} - 1)\alpha_{i_0} + \sum_{j \neq i_0} s_j \right\} = \max_{i=1, \dots, N} \left\{ (s_i - 1)\alpha_i + \sum_{j \neq i} s_j \right\} \quad (32)$$

证明: 在每个子矩阵 $\mathbf{A}_i$ 中, 选择指标集为 S_i 的列向量组成的子矩阵 $\mathbf{A}_{S_i}$ 。考虑 Grassmannian 矩阵

$$\begin{aligned} \Phi_S &= \mathbf{A}_S^T \mathbf{A}_S = \begin{pmatrix} \mathbf{A}_{S_1}^T \\ \vdots \\ \mathbf{A}_{S_N}^T \end{pmatrix} \begin{pmatrix} \mathbf{A}_{S_1} & \dots & \mathbf{A}_{S_N} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{A}_{S_1}^T \mathbf{A}_{S_1} & \mathbf{A}_{S_1}^T \mathbf{A}_{S_2} & \dots & \mathbf{A}_{S_1}^T \mathbf{A}_{S_N} \\ \mathbf{A}_{S_2}^T \mathbf{A}_{S_1} & \mathbf{A}_{S_2}^T \mathbf{A}_{S_2} & \dots & \mathbf{A}_{S_2}^T \mathbf{A}_{S_N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{S_N}^T \mathbf{A}_{S_1} & \mathbf{A}_{S_N}^T \mathbf{A}_{S_2} & \dots & \mathbf{A}_{S_N}^T \mathbf{A}_{S_N} \end{pmatrix} \end{aligned} \quad (33)$$

其中对角线上的块矩阵 $\mathbf{A}_{S_i}^T \mathbf{A}_{S_i}$ 具备以下形式

$$\begin{pmatrix} 1 & (\mathbf{a}_1^i)^T \mathbf{a}_2^i & \dots & (\mathbf{a}_1^i)^T \mathbf{a}_{s_i}^i \\ (\mathbf{a}_2^i)^T \mathbf{a}_1^i & 1 & \dots & (\mathbf{a}_2^i)^T \mathbf{a}_{s_i}^i \\ \vdots & \vdots & \ddots & \vdots \\ (\mathbf{a}_{s_i}^i)^T \mathbf{a}_1^i & (\mathbf{a}_{s_i}^i)^T \mathbf{a}_2^i & \dots & 1 \end{pmatrix}$$

其中, $\mathbf{a}_1^i, \dots, \mathbf{a}_{s_i}^i$ 是第 i 块中的列。利用 Gershgorin 圆盘定理, Φ_S 的最小特征值满足

$$\lambda_{\min} \geq 1 - \max_{i \neq j} \sum |\Phi_{ij}| \quad (34)$$

其中, Φ_{ij} 为 Φ_S 中的元素, 且 Φ_S 中对角线元素满足 $|\Phi_{ij}| \leq \mu^{i,j} = \alpha_i \mu$, 非对角线元素满足 $|\Phi_{ij}| \leq \mu$ 。因此

$$1 - \max_{i=1, \dots, N} ((s_i - 1)\alpha_i + (s - s_i))\mu \leq \lambda(A_S^T A_S) \leq 1 + \max_{i=1, \dots, N} ((s_i - 1)\alpha_i + (s - s_i))\mu \quad (35)$$

证毕

引理 2. 设矩阵 $A = [A_1, \dots, A_N]$ 的每一个子矩阵 A_i 的互相关参数是 α_i , 假设 1 成立, 则

$$A_S^T A_S \geq \hat{\gamma} I \quad (36)$$

其中,

$$\begin{aligned} \hat{\gamma} &= 1 - \left((s_I - 1)\alpha_I + \sum_{i \neq I} s_i \right) \mu \\ &= 1 - \max_{j=1, \dots, N} \left\{ (s_j - 1)\alpha_j + \sum_{i \neq j} s_i \right\} \mu \end{aligned} \quad (37)$$

定理 2. 设假设条件 1 和 2 均成立, 若分片稀疏信号满足

$$\begin{cases} x_{\min} \geq \max \left\{ \frac{4\sigma}{\sqrt{\hat{\gamma}}}, \frac{8\sigma(2 + \log_{10} s_{\max}) \left(\max_{j \in I} \|A_j\|_2 \right)}{\hat{\gamma} \hat{\eta}} \right\} \\ \text{且} \\ x_{\min} \geq \frac{2\sigma}{\sqrt{\hat{\gamma}}} \left(\sqrt{m + 2\sqrt{m \log_{10} m}} + \sqrt{\log_{10} s} \right) \end{cases} \quad (38)$$

其中, $s_{\max}$ 为 $s_i = |S_i|$ 的最大值;

$$\begin{aligned} \hat{\gamma} &= 1 - \left((s_I - 1)\alpha_I + \sum_{i \neq I} s_i \right) \mu \\ \hat{\eta} &= 1 - \frac{\sum_{i=1}^N \frac{\mu s_i}{1 + \mu s_i - (s_i - 1)\alpha_i \mu} - \frac{(1 - \alpha_Z) s_Z \mu}{1 + \mu s_Z - (s_Z - 1)\alpha_Z \mu}}{1 - \sum_{i=1}^N \frac{\mu s_i}{1 + \mu s_i - (s_i - 1)\alpha_i \mu}} \end{aligned} \quad (39)$$

则 aP_ISS 在终止准则 $\|r(t)\|_2 \leq \sigma \sqrt{m + 2\sqrt{m \log_{10} m}}$ 下以大于或等于 $1 - O(1/m)$ 的概率选出真实支集 S ($\hat{\eta}$ 的证明参见文献[23])。

3 数值实验

3.1 实验设计

本文的散乱数据曲面拟合问题, 对应于分片稀疏模型, 矩阵 $A = [A_1, \dots, A_N]$ 是由 ϕ_j^i 定义的观测矩阵 $A_i(s, k) = \phi \left(\frac{2^i x_s}{h} - k \right)$; $k \in K$, 带噪音的观测数

据 $\{f_i\}$ 是用向量 b^σ 表示。此部分通过给定的散乱点集 $\{x_k, y_k\}$ 和其相对应的观测数据 $\{f(x_k, y_k)\}$ 拟合曲面并得到其稀疏表示 $g(x, y)$ 。其逼近函数是由张量积 B 样条基函数 $\phi_j^i(x, y) (i=1, \dots, N; j=1, \dots, n_i)$ 表示, 其中 N 为多层 B 样条的层数, 基函数的总数为 $n = n_1 + \dots + n_N$, 于是待拟合曲面为

$$g(x, y) = \sum_{i=1}^N \sum_{j=1}^{n_i} c_j^i \phi_j^i(x, y) \quad (40)$$

考虑到 N 层 B 样条基函数之间是相关的, 于是此问题转变为了一个求解分片稀疏向量(曲面的系数向量), 即

$$c = (c_1^1, \dots, c_{n_1}^1, c_1^2, \dots, c_{n_2}^2, \dots, c_1^N, \dots, c_{n_N}^N)^T \quad (41)$$

的分片稀疏逼近问题。

本文通过对一个函数 f 添加高斯噪音来人工生成一个带噪音的数据集 $\{(x_k, y_k), f(x_k, y_k)\} : k = 1, \dots, 900\}$, 即

$$f = f(x_k, y_k) + \epsilon_k, \quad k = 1, \dots, 900 \quad (42)$$

其中, $\{(x_k, y_k)\}_{k=1}^{900} \subseteq \Omega = [-1, 1] \times [-1, 1]$ 是随机数据集, 高斯噪音 $\epsilon \in N(0, \sigma^2 I_n)$ 。测试层数为 $N=4$, 且测试的观测函数为

$$\begin{cases} f_1(x, y) = 0.75 \exp(-(9x-2)^2/4 - (9y-2)^2/4) + \\ \quad 0.75 \exp(-(9x+1)^2/49 - (9y+1)^2/10) + \\ \quad 0.5 \exp(-(9x-7)^2/4 - (9y-3)^2/4) - \\ \quad 0.2 \exp(-(9x-4)^2 - (9y-7)^2) \\ f_2(x, y) = \exp((-81(x-0.5)^2 - 81(y-0.5)^2)/4)/3 \\ f_3(x, y) = (-20.25(x-0.5)^2 + (y-0.5)^2)/3 \end{cases} \quad (43)$$

对比以下算法求解散乱数据曲面拟合问题的稀疏表示的数值效果:

- (1) 逆尺度空间算法(aISS)。
- (2) 带有删除机制的分片逆尺度空间算法(P_ISS), 其中删除机制中参数 $C=0.3$ 。
- (3) 多块 LASSO (MLASSO)

$$\min_x \sum_{i=1}^4 \lambda_i \|x_i\|_1 + \frac{1}{2} \left\| \sum_{i=1}^4 A_i x_i - b^\sigma \right\|_2^2 \quad (44)$$

利用 Jacobi 型的逼近交替方向乘子算法(alternating direction method of multipliers, ADMM)^[28] 求解 MLASSO, 其中正则化参数的值取为 $\lambda_1 = 0.3$, $\lambda_2 = 0.15$, $\lambda_3 = 0.2$, $\lambda_4 = 0.3$ 。

- (4) 本文提出的自适应分片逆尺度空间算法

(aP_ISS)。通过均方根误差(root-mean-square, RMS)、分片稀疏度及算法运行时间(CPU)进行 4 种算法比较, RMS 为

$$RMS = \frac{\sqrt{\sum_{k=1}^{M_1} \sum_{l=1}^{M_2} (f(x_k, y_l) - g(x_k, y_l))^2}}{M_1 M_2} \quad (45)$$

其中,

$$x_k = -1 + \frac{2k}{M_1 - 1} (k=0, 1, \dots, M_1 - 1) \quad (46)$$

$$y_l = -1 + \frac{2l}{M_2 - 1} (l=0, 1, \dots, M_2 - 1), M_1 = M_2 = 50$$

3 个噪音分别取值为 $\sigma_1 = 0.25, \sigma_2 = 0.1, \sigma_3 = 0.01$ 。

3.2 实验结果分析

从表 1~3 的数值结果和图 1~3 的拟合效果可以发现, 若不考虑分片稀疏性的 a_ISS 算法, 本文算法(aP_ISS)较其他 2 种算法在近似散乱点集上优势较明显, 同时在保护分片稀疏度上效果更好, 即每一层支集均不为空集且稀疏。相较于分片稀疏性的 MLASSO 算法, aP_ISS 算法在运行效率上有明显的优势, MLASSO 模型存在事先人工输入和调节多个正则化参数的缺陷, 本文算法的自适应性保证了算法可以更快地运行, 且 MLASSO 算法近似的曲面效果远不如本文算法。P_ISS 算法的主要缺陷在于删除机制中参数 C 的选取依赖数据、噪音等因素, 需要花费大量时间, 而本文算法不仅避免了参数的选取而且能够得到和 P_ISS 算法同等的近似效果。

另从表 1~3 的结果还可以观察到, 分片稀疏逼近的目的是为了在每一层上都有非零元素表示, 且是稀疏表示, 本文算法除了进一步改进已有的分片算法 P_ISS, 更重要的是为了通过“分片符号一致性”来体现分片稀疏的思想。实验结果表明本文的算法不仅在拟合效果上更好, 并且达到了分片稀疏表示的效果。

表 1 观测函数 f_1 对应曲面拟合的数值结果

算法	RMS	分片稀疏度	CPU
a_ISS	0.388	(3+0+0+0) 3	0.009
P_ISS	0.394	(3+2+2+2) 9	0.015
MLASSO	0.408	(3+22+46+9) 80	0.126
本文	0.394	(3+2+2+2) 9	0.015

表 2 观测函数 f_2 对应曲面拟合的数值结果

算法	RMS	分片稀疏度	CPU
a_ISS	0.112	(1+0+0+0) 1	0.008
P_ISS	0.118	(0+1+1+0) 2	0.005
MLASSO	0.118	(1+14+9+0) 24	0.129
本文	0.118	(0+1+1+0) 2	0.006

表 3 观测函数 f_3 对应曲面拟合的数值结果

算法	RMS	分片稀疏度	CPU
a_ISS	0.108	(7+21+13+0) 41	0.054
P_ISS	0.107	(11+14+15+13) 53	0.059
MLASSO	0.104	(3+28+2+0) 33	0.124
本文	0.107	(11+14+15+13) 53	0.059

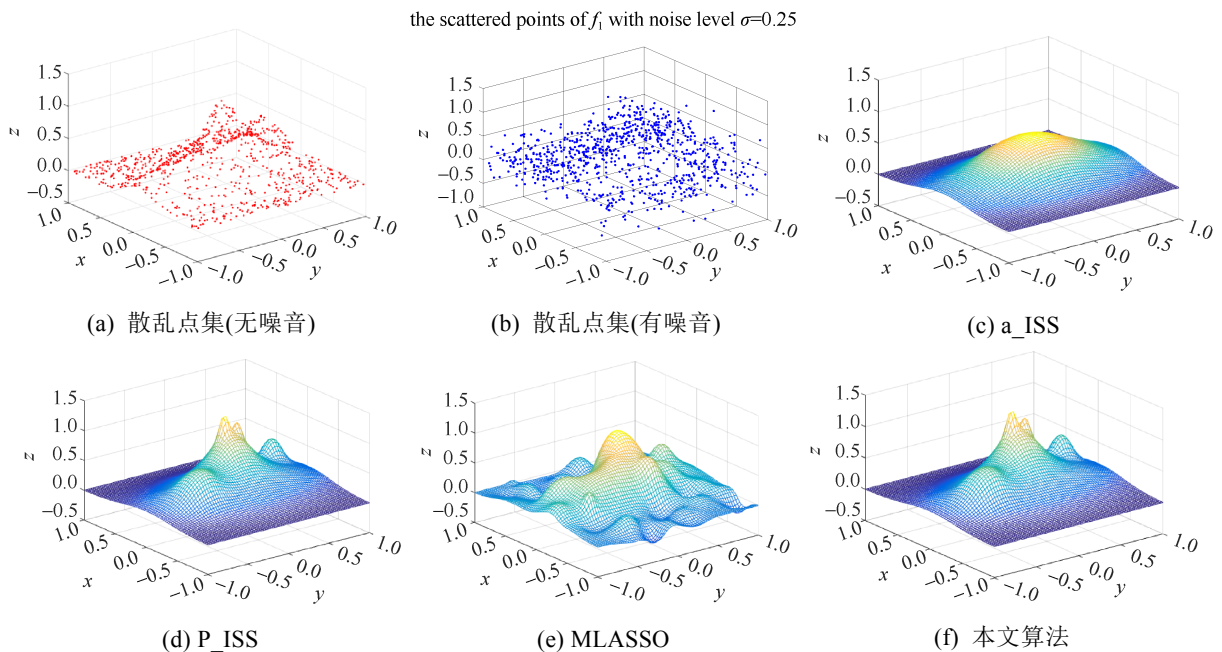
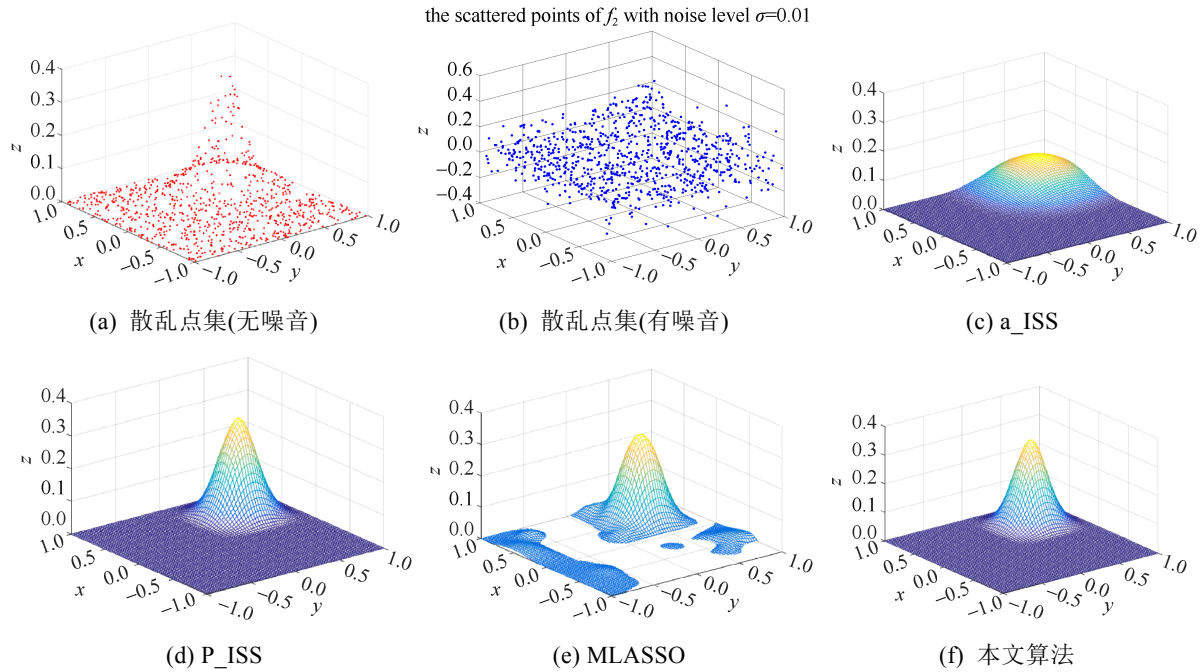
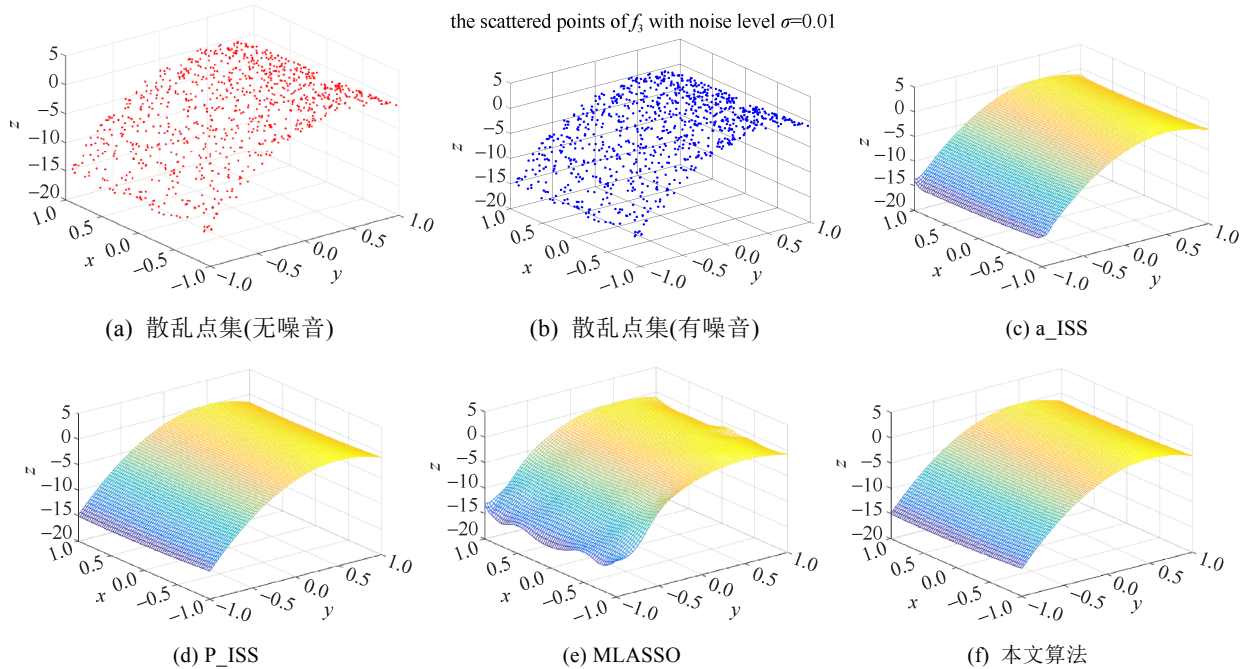


图 1 观测函数 f_1 对应 4 种算法拟合出的近似曲面的对比图

图 2 观测函数 f_2 对应 4 种算法拟合出的近似曲面对比图图 3 观测函数 f_3 对应 4 种算法拟合出的近似曲面对比图

4 结 束 语

本文利用 aP_ISS 处理散乱点曲面拟合问题, 通过引入分片稀疏性, 可以有效地保护在 PSI 空间中多层支集的分片稀疏性; 进一步通过对分片逆尺度空间系统的分片符号一致性的考虑, 避免了分片算法中参数的选取问题。实验结果表明, 该算法近似散乱点曲面效果较好, 验证了利用稀疏优化算法处理散乱点曲面拟合问题的优势。

分片稀疏思想应用到散乱数据拟合等问题仍然欠缺较为完善的研究, 通过对稀疏优化算法的自适应分片处理来求解曲面拟合问题, 将是未来研究的重点。

参 考 文 献

- [1] LEE S, WOLBERG G, SHIN S Y. Scattered data interpolation with multilevel B-splines[J]. IEEE Transactions on Visualization and Computer Graphics, 1997, 3(3): 228-244.

- [2] CASTANO D, KUNOTH A. Multilevel regularization of wavelet based fitting of scattered data some experiments[J]. Numerical Algorithms, 2005, 39(1-3): 81-96.
- [3] JOHNSON M J, SHEN Z, XU Y H. Scattered data reconstruction by regularization in B-spline and associated wavelet spaces[J]. Journal of Approximation Theory, 2009, 159(2): 197-223.
- [4] HUANG J Z. Structured sparsity: theorems, algorithms and applications[D]. New Jersey: Rutgers University, 2011.
- [5] ELAD M, BRUCKSTEIN A M. A generalized uncertainty principle and sparse representation in pairs of bases[J]. IEEE Transactions on Information Theory, 2002, 48(9): 2558-2567.
- [6] ELDAR Y C, MISHALI M. Block sparsity and sampling over a union of subspaces[C]//2009 16th International Conference on Digital Signal Processing. New York: IEEE Press, 2009: 1-8.
- [7] ELDAR Y C, KUPPINGER P, BOLCSKEI H. Block-sparse signals: uncertainty relations and efficient recovery[J]. IEEE Transactions on Signal Processing, 2010, 58(6): 3042-3054.
- [8] DUARTE M F, ELDAR Y C. Structured compressed sensing: from theory to applications[J]. IEEE Transactions on Signal Processing, 2011, 59(9): 4053-4085.
- [9] ELHAMIFAR E, VIDAL R. Block-sparse recovery via convex optimization[J]. IEEE Transactions on Signal Processing, 2012, 60(8): 4094-4107.
- [10] BACH F. Consistency of the group LASSO and multiple kernel learning[EB/OL]. [2019-05-23]. <https://arxiv.org/abs/0707.3390v1>.
- [11] FRIEDMAN J H, HASTIE T, TIBSHIRANI R. A note on the group lasso and a sparse group lasso[EB/OL]. [2019-06-04]. <https://arxiv.xilesou.top/abs/1001.0736>.
- [12] HUANG J Z, ZHANG T. The benefit of group sparsity[J]. The Annals of Statistics, 2010, 38(4): 1978-2004.
- [13] JACOB L, OBOZINSKI G, VERT J P. Group lasso with overlap and graph lasso[C]//Proceedings of the 26th Annual International Conference on Machine Learning - ICML '09. New York: ACM Press, 2009: 433-440.
- [14] MEIER L, VAN DE GEER S, BÜHLMANN P. The group lasso for logistic regression[J]. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2008, 70(1): 53-71.
- [15] ROTH V, FISCHER B. The Group-LASSO for generalized linear models: uniqueness of solutions and efficient algorithms[C]//Proceedings of the 25th International Conference on Machine Learning. New York: Association for Computing Machinery, 2008: 848-855.
- [16] WANG H S, LENG C L. A note on adaptive group lasso[J]. Computational Statistics and Data Analysis, 2008, 52(12): 5277-5286.
- [17] YUAN M, LIN Y. Model selection and estimation in regression with grouped variables[J]. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2006, 68(1): 49-67.
- [18] BARANIUK R G, CEVHER V, DUARTE M F, et al. Model-based compressive sensing[J]. IEEE Transactions on Information Theory, 2010, 56(4): 1982-2001.
- [19] 刘翼鹏. 基于凸优化的参数化稀疏估计理论及其应用[D]. 成都: 电子科技大学, 2011.
- [20] HAO Y X, LI C J, WANG R H. Sparse approximate solution of fitting surface to scattered points by MLASSO model[J]. Science China Mathematics, 2018, 61(7): 1319-1336.
- [21] 钟铁君. 分片稀疏恢复理论及算法[D]. 大连: 大连理工大学, 2018.
- [22] OSHER S, RUAN F, XIONG J C, et al. Sparse recovery via differential inclusions[J]. Applied and Computational Harmonic Analysis, 2016, 41(2): 436-469.
- [23] LI C J, ZHONG Y J. Piecewise sparse recovery in unions of bases[EB/OL]. [2019-06-04]. <https://arxiv.xilesou.top/abs/1903.01208>.
- [24] DONOHO D L, HUO X. Uncertainty principles and ideal atomic decomposition[J]. IEEE Transactions on Information Theory, 2001, 47(7): 2845-2862.
- [25] BURGER M, MÖLLER M, BENNING M, et al. An adaptive inverse scale space method for compressed sensing[J]. Mathematics of Computation, 2013, 82(281): 269-299.
- [26] NEEDELL D, TROPP J A. CoSaMP: iterative signal recovery from incomplete and inaccurate samples[J]. Applied and Computational Harmonic Analysis, 2009, 26(3): 301-321.
- [27] DONOHO D L, TSAIG Y, DRORI I, et al. Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit[J]. IEEE Transactions on Information Theory, 2012, 58(2): 1094-1121.
- [28] DENG W, LAI M J, PENG Z M, et al. Parallel multi-block ADMM with $o(1/k)$ convergence[J]. Journal of Scientific Computing, 2017, 71(2): 712-736.