约束优化模型的稳定性
用于结构化信号恢复

Yijun Zhong、Yi Shen

arXiv:2601.04849v1 [cs.IT] 8 Jan 2026

摘要

从测量值中恢复未知但具有结构的信号是一项具有挑战性的问题，在图像复原、无线通信和信号处理等领域具有重要应用。本文研究了由信号结构的先验知识所引发的内在问题，例如对信号恢复模型至关重要的稀疏性。我们研究了三种能够有效应对这一挑战的约束优化模型，每种模型都利用不同形式的结构先验来正则化解空间。理论分析表明，这些模型具有抗噪声鲁棒性，并且相对于调节参数保持稳定；这是一项对实际应用至关重要的性质，因为参数选择往往并非易事。通过提供理论基础，我们的工作支持在测量误差和模型不确定性不可避免的场景中实际应用这些模型。此外，在温和条件下，我们建立了样本复杂度与失配误差之间的权衡关系。

索引词

最小二乘法、广义 Lasso、稀疏相位恢复、非线性逼近、高斯宽度。

I. 引言

现代信息处理和机器学习不断面临如何有效处理结构化信号这一挑战，无论这些信号是数据、图像还是声音。本文主要关注如何根据相对少量的含噪测量值 $\boldsymbol{y} \in \mathbb{R}^m$ 和测量矩阵$A \in \mathbb{R}^{m\times n}$ （其中 $m < n$），恢复向量 $\boldsymbol{x}^* \in \mathbb{R}^n$ 。由于这个问题通常适定性不良，因此只有当其具有某种结构时，才能恢复 $\boldsymbol{x}^*$ 。我们假设已知关于有界结构诱导函数 f 在 $\boldsymbol{x}^*$ 处的先验信息，据此可以定义一个可行集

$$\mathcal{K} = \{\boldsymbol{x} \in \mathbb{R}^n : f(\boldsymbol{x}) \leq \eta\},$$

其中 $\eta = f(\boldsymbol{x}^*)$称为最优调节参数。可行集 $\mathcal{K}$ 刻画了 $\boldsymbol{x}^*$的结构，涵盖多种类型，例如所有 s-稀疏信号的集合、经过适当缩放的 l_1 球，以及

本研究受国家自然科学基金委员会资助，项目编号为 12371101 和 11901529。（通讯作者：*Yi Shen*）Yijun Zhong 任职于中国杭州浙江理工大学数学系，邮编 310018。（电子邮件：zhongyi-jun@zstu.edu.cn）。

Yi Shen 任职于中国杭州浙江理工大学数学系，邮编 310018。（电子邮箱：yshen@zstu.edu.cn）。

星形集合 [1]。我们考虑以下三种以可行集 $\mathcal{K}$ 为约束的结构信号恢复优化模型:

- 约束最小二乘

$$\min_{\boldsymbol{x}\in\mathcal{K}} \frac{1}{2}\|\boldsymbol{y}-A\boldsymbol{x}\|_2^2. \tag{1}$$

- 约束最小绝对偏差

$$\min_{\boldsymbol{x}\in\mathcal{K}} \|\boldsymbol{y}-A\boldsymbol{x}\|_1. \tag{2}$$

- 约束非线性最小二乘

$$\min_{\boldsymbol{x}\in\mathcal{K}} \frac{1}{2}\left\|\boldsymbol{y}-|A\boldsymbol{x}|\right\|_2^2. \tag{3}$$

A. 研究动机

可行集 $\mathcal{K}$ 界定了可能解的范围，使各种应用中的问题求解能够更加聚焦、高效。然而，确定调节参数 η的最优取值并非总是轻而易举。交叉验证是一种广泛认可的近似确定合适取值的方法。一个自然的问题是，约束优化模型的性能是否会受到调节参数具体选择的影响。关于稀疏信号恢复模型对参数选择的敏感性的研究，可见 [2]–[7] 及其中的参考文献。特别是， [4], [5] 中的研究评估了三种 LASSO 变体在其控制参数方面的极小极大阶最优恢复结果。在 [4],中，作者研究了近端去噪问题，并刻画了噪声尺度趋于零时风险的渐近奇异性。具体而言， [4, 定理 2.1] 考察了三种不同的情形： $\|\boldsymbol{x}^*\|_1<\eta$、 $\|\boldsymbol{x}^*\|_1=\eta$和 $\|\boldsymbol{x}^*\|_1>\eta$。随后， [5, 定理 V.1（渐近奇异性）] 将这一分析扩展到了满足 $\|\boldsymbol{x}^*\|_1\neq\eta$ 的约束 Lasso。受这些关于调节参数 η的渐近结果启发，我们针对结构信号恢复问题推导了非渐近误差界，其中 $f(\boldsymbol{x}^*)\neq\eta$。

B. 主要贡献

我们证明了 以下两种情况下约束优化模型 (1)、(2) 和 (3) 的稳定性

$$f(\boldsymbol{x}^*)<\eta \quad \text{and} \quad f(\boldsymbol{x}^*)\geq\eta.$$

条件 $f(\boldsymbol{x}^*)\neq\eta$ 可以从以下直觉来理解：近似误差受约束集 [4] 的有效维数控制。一个关键发现是，线性估计量和非线性估计量都能抵御调节参数的不准确，表明即使无法精确获知最优超参数，它们仍具有鲁棒性。在实际应用中，这种稳定性尤为重要，因为计算限制或先验知识有限可能使精确调节参数变得困难。另一方面，我们的主要结果量化了样本复杂度（测量值的数量）与失配误差（ $f(\boldsymbol{x}^*)$ 与可行集 $\mathcal{K}$ 的调节参数 η 之间的距离）之间的权衡。这种更精细的权衡从理论上阐明了测量限制与调节参数之间的相互作用，并为压缩感知、高维统计和逆问题中的算法设计提供了实用指南。

C. 路线图

本文结构安排如下。在本节余下部分，我们将介绍三种涉及线性测量和非线性测量的典型模型。第二节回顾与 (1)、(2) 和 (3) 的稳定性分析相关的工作，以及求解这些问题的算法。第三节简要回顾高斯宽度及其在可行集中的应用。第四节确立本文讨论的三种模型的稳定性。部分引理和定理的证明见补充材料。

D. 符号说明

对于任意给定的向量 $\boldsymbol{x}=(x_1,x_2,\ldots,x_n)^T$，其“$\ell_0$–范数”记为 $\|\boldsymbol{x}\|_0$，是其非零元素的数量。若向量满足 $\|\boldsymbol{x}\|_0\le s$，则称其为 s-稀疏向量。对于任意 $p>0$，定义 $\|\boldsymbol{x}\|_p:=(\sum_{i=1}^n|x_i|^p)^{1/p}$。对于任意给定的正整数 m，记为 $[m]=\{1,\ldots,m\}$。给定指标集 $\Omega\subset[m]$ 和向量 $\boldsymbol{x}$，令 $\boldsymbol{x}_\Omega$ 表示一个向量，其第 i 个元素等于 $\boldsymbol{x}$ 的第 i 个元素（ i 属于 Ω 时），否则为零。类似地，令 A_Ω表示一个矩阵，其第 i 行等于 A 的第 i 行（ i 属于 Ω 时），否则为零向量。字母C、 C_1和 C_2 通常视为常数，但它们的取值会因论文的不同部分而异。我们不显式写出 $a\le Cb$，而写作 $a\lesssim b$；不写 $a\ge Cb$，而写作 $a\gtrsim b$，其中分别有 $C>1$。我们使用记号 $a\approx b$ 表示存在常数 $C_1>0$ 和 $C_2>0$，使得 $C_1b\le a\le C_2b$.

令 $\mathcal{N}(\boldsymbol{0},I_n)$ 表示 $\mathbb{R}^n$ 中均值为零、协方差矩阵为单位矩阵的多元正态分布。我们假设测量矩阵 $A\in\mathbb{R}^{m\times n}$ 是一个高斯随机矩阵，其各行 $\boldsymbol{a}_i\sim\mathcal{N}(\boldsymbol{0},I_n)$ 相互独立。单位欧几里得球面记为

$$\mathcal{S}^{n-1}=\{\boldsymbol{x}\in\mathbb{R}^n:\ \|\boldsymbol{x}\|_2=1\},$$

$\mathbb{R}^n$ 中的单位欧几里得球记为

$$\mathcal{B}_2^n=\{\boldsymbol{x}\in\mathbb{R}^n:\ \|\boldsymbol{x}\|_2\le 1\}.$$

$\boldsymbol{x}^*$ 到集合 $\mathcal{K}$ 的欧几里得投影记为

$$\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)\in\arg\min_{\boldsymbol{x}\in\mathcal{K}}\|\boldsymbol{x}^*-\boldsymbol{x}\|_2^2.$$

注意，我们不要求 $\mathcal{K}$ 是凸集，因此投影算子 $\boldsymbol{P}_{\mathcal{K}}$ 可能不唯一。对于给定的随机过程 $(X_{\boldsymbol{x}})_{\boldsymbol{x}\in\mathcal{T}}$，将指标集 $\mathcal{T}$ 上的典范度量记为 $d(\boldsymbol{x},\boldsymbol{y})=(\mathbb{E}(X_{\boldsymbol{x}}-X_{\boldsymbol{y}})^2)^{\frac{1}{2}}$。对于任意给定的$\mathcal{T}$，集合 $\mathcal{T}$的直径记为

$$\operatorname{diam}(\mathcal{T})=\sup\{\|\boldsymbol{x}-\boldsymbol{y}\|_2:\boldsymbol{x},\boldsymbol{y}\in\mathcal{T}\}$$

$\mathcal{T}$的半径记为 $\operatorname{rad}(\mathcal{T})=\sup\ \|\boldsymbol{x}\|_2$。我们还采用 $\mathcal{K}$ 的闵可夫斯基泛函这一概念，其定义$_{\boldsymbol{x}\in\mathcal{T}}$为

$$\|\boldsymbol{x}\|_{\mathcal{K}}=\inf\{\lambda>0:\ \lambda^{-1}\boldsymbol{x}\in\mathcal{K}\}$$

对于给定的向量 $\boldsymbol{x}\in\mathbb{R}^n$。

II. 相关工作

过去几年中，文献中涌现了大量关于具有结构性质的信号恢复模型稳定性的研究，包括求解此类模型的算法的收敛性分析。这些研究深受相对新兴的压缩感知领域的启发 [8], ，其理论基础源自几何泛函分析 [9] 和凸积分几何 [10]中的经典结果。关于稀疏恢复问题的相关研究大致可分为两类：线性方法和非线性方法。为简洁起见，我们聚焦于实值情形。

A. 线性估计

我们从线性逆问题开始:

$$\boldsymbol{y} = A\boldsymbol{x}^* + \boldsymbol{e}$$

其中， $\boldsymbol{e}$ 表示与测量矩阵相互独立的噪声向量。第一个恢复模型是在结构约束（1）下的标准最小二乘法。信号的不同结构信息$\boldsymbol{x}$ 由 $\mathcal{K}$ 表示，由此产生了大量得到广泛研究的问题。例如，如果 f 是由稀疏性诱导的“l_0 范数”，那么模型（1）就是子集选择。如果 f 是 l_1 范数，那么在压缩感知文献中，模型（1）被称为基追踪（BP）；在统计回归领域，则称为最小绝对收缩和选择算子（套索回归） [12]。在 [2], 中，Chatterjee 研究了凸约束下的最小二乘问题，这是（1）的一个特例，并说明对于非奇异设计矩阵，当 f（$\boldsymbol{x}^*$）$= \eta$时，预测误差会非常小。大量研究致力于求解稀疏约束优化（1）的算法，例如 [13]–[19] 及其中的许多参考文献。如果 f 是具有 $0 < p \leq 1$的非凸稀疏性诱导 l_p 范数，那么人们使用了两类迭代算法，通过求解 l_p约束的最小二乘问题来近似 $\boldsymbol{x}^*$:

$$\min_{\|\boldsymbol{x}\|_p \leq \eta} \frac{1}{2}\|\boldsymbol{y} - A\boldsymbol{x}\|_2^2. \tag{4}$$

一种是投影梯度下降算法 [16], [19]。在 $\|\boldsymbol{x}^*\|_p \geq \eta$ 的情况下，已证明投影梯度下降算法在 受限等距性质 (RIP) 条件[19]下对调节参数 η 具有稳定性。随后， [16] 证明，在情况 $\|\boldsymbol{x}^*\|_p \leq \eta$ 和情况 $\|\boldsymbol{x}^*\|_p \geq \eta$下，采用固定学习率的投影梯度下降算法对调节参数均具有稳定性。[16], [19], [20] 中关于迭代算法的稳定性分析与定理 IV.1 一致。

对于具有重尾噪声或离群值的线性模型，最小二乘法 (1) 在信号恢复方面往往表现不佳，因此需要采用最小绝对偏差 (LAD) 模型 (2)。如果 f 选取为 l_1 范数，且调节参数满足 $\|\boldsymbol{x}^*\|_1 \leq \eta$，那么模型 (2)

$$\min_{\|\boldsymbol{x}\|_1 \leq \eta} \|\boldsymbol{y} - A\boldsymbol{x}\|_1 \tag{5}$$

对于稠密噪声和对抗性稀疏噪声，均显示出很强的鲁棒性 [7]。然而， $\|\boldsymbol{x}^*\|_1 > \eta$的情况在 [7]中未被讨论。我们将在 [7, 定理 1.1] 之后关于推论 IV.3 的讨论中补充这一点。通过求解约束最小绝对偏差来削弱对抗性稀疏噪声的方法，进一步扩展到了低秩矩阵恢复

问题 [21] 和相位恢复问题 [22]。如果在模型 (2) 中将 f 选为“l_0–范数”，那么文献 [23], [24] 提出了投影次梯度下降方法，可直接求解 (2)。统计学中一个著名的模型是 l_1 惩罚最小绝对偏差。文献 [25]讨论了变量选择一致性的渐近性质。文献 [26], [27]讨论了 l_1 惩罚最小绝对偏差估计量的一致性。文献 [28]得到了 l_1 惩罚最小绝对偏差的近似预言机性能。

B. 非线性估计—相位恢复

我们针对线性模型的分析方法可以扩展到其他相关的恢复问题，例如相位恢复、低秩矩阵恢复和盲反卷积。我们通过验证一个知名相位恢复模型的稳定性说明了这一点。从数学上看，相位恢复模型所描述的测量值为

$$\boldsymbol{y} = |A\boldsymbol{x}^*| + \boldsymbol{e},$$

其中， $\boldsymbol{e}$ 表示与测量矩阵独立的噪声。一种恢复 $\boldsymbol{x}^*$ 的方法是求解模型 (3)。尽管这一二次测量方案具有非线性特征，但事实证明，它能够有效捕获仅强度观测，同时保持数学上的可处理性。模型 (3) 在叠层成像等实际领域中表现良好，例如芯片成像 [29]。与基于强度的模型相比，最小化基于振幅的模型 (3) 的算法通常计算效率更高 [29]–[31]。Huang 和 Xu 在 [32] 证明，当调节参数取最优值，即 $\|\boldsymbol{x}^*\|_1 = \eta$ 时，模型 (3) 对噪声具有鲁棒性。此外， [33, 定理 1.2] 中的结果通过采用不同的工具，将 [32, 定理 1.5] 中针对实信号的误差界推广到了复信号。文献 [34]提出了两种利用生成式先验求解相位恢复问题的算法。文献 [29]提出了投影 *Wirtinger* 流（PWF）来求解模型 (3)。只要将 PWF 初始化在未知信号 $\boldsymbol{x}^*$的适当邻域内，就能提供收敛保证。为在满足最优条件 $\|\boldsymbol{x}\|_0 \leq \eta$ 且 $\|\boldsymbol{x}^*\|_0 = \eta$的情况下求解模型 (3)，人们设计了许多算法，并在稀疏相位恢复问题上取得了良好效果，例如迭代硬阈值 [35], 截断幅度流 [30], CoPRAM [36]。现有研究 [30-34] 侧重于模型的稳定性分析或算法的收敛性分析，且都假设 η 为最优参数。这一疏漏意味着， η 偏离最优值对模型稳定性和算法收敛性的潜在影响仍未得到探索，使我们对这些分析框架的理解存在空白。

III. 预备知识

本节回顾高斯宽度的一些性质。高斯宽度广泛用于从独立随机线性测量中恢复结构化信号的凸恢复问题中 [37], [38], ，也用于统计学和信号处理，以量化一组数据点的变异性或离散程度 [16], [39], [40]。

定义 III.1 (高斯宽度)。 集合 $\mathcal{K} \subset \mathbb{R}^n$ 的高斯宽度定义如下：

$$\omega(\mathcal{K}) = \mathbb{E}_{\boldsymbol{g}} \left[\sup_{\boldsymbol{z} \in \mathcal{K}} \langle \boldsymbol{g}, \boldsymbol{z} \rangle \right],$$

其中，期望是对标准高斯随机向量 $\boldsymbol{g} \sim \mathcal{N}(0, I_n)$取的。

高斯宽度与源自几何泛函分析和凸几何的平均宽度相似。对于给定的集合 $\mathcal{K}$，高斯宽度 $\omega(\mathcal{K})$ 可以由 $\mathcal{K}$ [40, 的覆盖数界定定理 3.11]。

定义 III.2（下降集与锥）。 *[16, 定义 1]* 函数 f 在点 $\boldsymbol{x}$处的下降集定义为

$$\mathcal{D}_f(\boldsymbol{x}) = \{\boldsymbol{h} : \ f(\boldsymbol{x}+\boldsymbol{h}) \leq f(\boldsymbol{x})\}.$$

下降锥定义为包含下降集的闭锥 $\mathcal{C}_f$（$\boldsymbol{x}$），即 $\mathcal{D}_f$（$\boldsymbol{x}$）$\subset \mathcal{C}_f$（$\boldsymbol{x}$）。切锥是下降集的锥包。

$\mathcal{C}_f(\boldsymbol{x})$ 的另一种定义可见于 [3], ；正如 [40], 所示，形式为 $\mathcal{T} = \mathcal{C}_f(\boldsymbol{x}) \cap \mathcal{S}^{n-1}$ 的单位球面子集的高斯平均宽度在结构化恢复中起着重要作用，而$\omega^2(\mathcal{C}_f(\boldsymbol{x}) \cap \mathcal{S}^{n-1})$ 是 $\mathcal{K}$ [39]维数的一种近似。

定义 III.3（相变函数）。 *[16, 定义 4]* 设 $\mathcal{C}_f$（$\boldsymbol{x}$）是 f 在 $\boldsymbol{x}$处的下降锥。令$\omega = \omega$（$\mathcal{C}_f$（$\boldsymbol{x}$）$\cap \mathcal{B}^n$），令 ϕ（t）$= \sqrt{2}^{\Gamma\left(\frac{t+1}{2}\right)\Gamma\left(\frac{t}{2}\right)} \approx \sqrt{t}$, 一则相变函数定义为

$$\mathcal{M}(f, \boldsymbol{x}, u) = \phi^{-1}(\omega + u) \approx (\omega + u)^2$$

其中， $u > 0$ 是控制成功概率的参数。

我们使用简写 $m_0 = \mathcal{M}(f, \ \boldsymbol{x}, \ u)$，并隐含其对 f、 $\boldsymbol{x}$、 u 的依赖关系。 [16] 中证明，对于凸 f，m_0 恰好是程序（1）以高概率成功恢复 $\boldsymbol{x}^*$ 所需的最小测量数。此外，我们定义了 $\mathcal{M}(f, \ \boldsymbol{x})$，它近似刻画了所需的最小样本数。

定义 III.4。 令 $\mathcal{C}_f(\boldsymbol{x})$ 为 f在 $\boldsymbol{x}$处的下降锥。令 $\omega = \omega(\mathcal{C}_f(\boldsymbol{x}) \cap \mathcal{S}^{n-1})$，则近似最小样本函数定义为 $\mathcal{M}(f$、$\boldsymbol{x})\ = \omega^2$.

为方便起见，我们将 $m_1 = \mathcal{M}(f, \ \boldsymbol{x})$ 记作该函数。关于最小样本数的其他类似设定，可参见 [29, 定义 3]。对（1）、（2）和（3）的估计依赖于可行集 $\mathcal{K}$ 的大小，因此我们基于高斯宽度提供上界，以量化一般子集 $\mathcal{K} \subset \mathbb{R}^n$ 的复杂度。该锥越小，函数 f 就越适合捕捉 $\boldsymbol{x}^*$的性质。高斯宽度是评估可行集 $\mathcal{K}$大小的有力工具，其优势在于对扰动具有稳健性：如果 $\mathcal{K}$ 略有增加，那么高斯宽度也只会发生边际变化 [1]。

定义 III.5. 一个函 数 f： $\mathbb{V} \to \mathbb{R}$ （其中 $\mathbb{V}$ 是一个向量空间）称为绝对齐次的，如果它满足 *s*

$$f(\alpha \boldsymbol{x}) = |\alpha| f(\boldsymbol{x}).$$

对于所有标量 $\alpha \in \mathbb{R}$ 和向量 $\boldsymbol{x} \in \mathbb{V}$。

例如，范数是绝对齐次的。在 $0 < f(\boldsymbol{x}^*) < \eta$ 的情况下，第四节所述主要理论结果中对 f 的绝对齐次性假设是必要的，因为它使得能够使用中间变量 $\boldsymbol{x}^\eta$。未来的工作可以研究，在不施加其他假设的情况下，是否能够得到（定理 IV.1 中的）(7) 的类似版本。

本节其余部分中，我们首先给出若干与随机过程相关的引理，随后刻画具有高斯宽度的高斯随机矩阵的近似各向同性。然后，我们给出关于随机矩阵在集合上的偏差的引理，从而为高斯随机矩阵建立双侧界。

引理 III.1. *[29, 定理 27]* 假设 $A \in \mathbb{R}^{m\times n}$ 是一个具有独立 $\boldsymbol{a}_i \sim \mathcal{N}(0, I_n)$ 行的高斯随机矩阵，则对于任意子集 $\mathcal{T} \subset \mathbb{R}^n$ 以及任意 $u \geq 0$, $\delta \in (0,1)$，事件

$$\left| \frac{\|A\boldsymbol{x}\|_2}{\phi(m)} - \|\boldsymbol{x}\|_2 \right| \leq \delta \|\boldsymbol{x}\|_2$$

对所有 $\boldsymbol{x} \in \mathcal{T}$ 以至少 $1 - 2\exp(-\frac{u^2}{2})$ 的概率成立，只要 $m \geq (\omega(\mathcal{T}) + u^2)$ _____.

引理 III.2. *[41, 推论 1.5]* 在引理 *III.1* 的假设下，对于任意 $u \geq 0$，事件

$$\sup_{\boldsymbol{x}\in\mathcal{T}\cap\mathcal{S}^{n-1}} \left| \|A\boldsymbol{x}\|_2 - \sqrt{m} \right| \leq C[\omega(\mathcal{T}\cap\mathcal{S}^{n-1}) + u],$$

以至少 $1 - 2\exp(-u^2)$ 的概率成立。

下一个引理给出了引理 III.1 和引理 III.2 的另一个版本，其由 m_0 定义于定义 III.3 中。

引理 III.3。 *[17, 引理 7。5]* 设 $\mathcal{C} = \mathcal{C}_f(\boldsymbol{x})$ 且 $m_0 = \mathcal{M}(f, \boldsymbol{x}, u)$，假设 $A \in \mathbb{R}^{m\times n}$ 是一个具有独立 $\boldsymbol{a}_i \sim \mathcal{N}(0, I_n)$ 行的高斯随机矩阵，则以下结论以至少 $1 - 2\exp(-u^2)$ 的概率成立：

$$\inf_{\boldsymbol{z}\in\mathcal{C}\cap\mathcal{S}^{n-1}} \|A\boldsymbol{z}\|_2^2 \geq (\phi(m) - \phi(m_0))^2.$$

引理 III.4。 *[16, 引理 23]* 定义 $\phi(t) = \sqrt{2}\frac{\Gamma(\frac{t+1}{2})}{\Gamma(\frac{t}{2})} \approx \sqrt{t}$ 如定义 *III.3* 中所示，则对于 $0 \leq m_0 \leq m$，我们有

$$\frac{\phi(m_0)}{\sqrt{m_0}} \leq \frac{\phi(m)}{\sqrt{m}}.$$

引理 III.5。 假设 $A \in \mathbb{R}^{m\times n}$ 是一个具有独立 $\boldsymbol{a}_i \sim \mathcal{N}(0, I_n)$ 行的高斯随机矩阵，则对于任意子集 $\mathcal{T} \subset \mathcal{B}_2^n$ 和任意 $u \geq 0$，事件

$$\sup_{\boldsymbol{x}\in\mathcal{T}, \boldsymbol{e}\in\mathbb{R}^m} |\langle \boldsymbol{x}, A^T\boldsymbol{e}\rangle| \leq \|\boldsymbol{e}\|_2[\omega(\mathcal{T}) + u]$$

以至少 $1 - 2\exp(-Cu^2)$ 的概率成立。

引理 III.5 的证明见附录。

Lemma III.6. 在引理 *III.1* 的条件下，假设 $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ 是一个固定向量，则 n

$$\min_{\Omega\subset[m],\ |\Omega|\geq m/2} \|A_\Omega \bar{\boldsymbol{x}}\|_2 \geq \frac{v_0}{2}\sqrt{m}\|\bar{\boldsymbol{x}}\|_2$$

以概率 $1-2\exp\left(-\frac{v_0^2}{8}m\right)$ 成立，且 $v_0=\frac{1}{32e}\sqrt{\frac{\pi}{2}}\left(1-\frac{1}{4\sqrt{\pi}}\right)\approx 0.0124$。

证明。 引理 III.6 的证明与 [42,引理 4.4]的证明类似。不失一般性，假设 $\|\bar{\boldsymbol{x}}\|_2=1$。并令 $\boldsymbol{z}=A\bar{\boldsymbol{x}}$、$\boldsymbol{z}_\Omega=A_\Omega\bar{\boldsymbol{x}}$，因此 $\boldsymbol{z}$ 的各个条目是高斯随机变量 $\boldsymbol{z}_i\sim\mathcal{N}(0,\ 1)$ 的独立实现。应用 [42, 引理 4.2]，令 $m\geq 1$ 和 [42,，通过设置 $t=\frac{v_0}{2}$，可得

$$\mathbb{P}\left[\min_{\Omega\subset[m],\ |\Omega|\geq m/2}\frac{1}{m}\|\boldsymbol{z}_\Omega\|_2^2\geq\frac{v_0^2}{4}\right]\geq 1-2\exp\left(-\frac{v_0^2}{8}m\right).$$

□

IV. 主要结果

本节介绍我们的主要结果。为便于分析，我们将可能取值的范围进行分类。每种情况都呈现出独特的性能特征，因此需要采用不同的估计方法。

A. 线性估计

定理 IV.1。 设 $\boldsymbol{x}^*$ 为 $\mathbb{R}^n$ 中的任意向量。对于任意 $u>0$，令 $m_0=\mathcal{M}(f、\ \boldsymbol{x}^*、\ u)$，其定义见定义 *III.3*。设 $A\in\mathbb{R}^{m\times n}$ 为具有独立 $\boldsymbol{a}_i\sim\mathcal{N}(0、\ I_n)$ 行的高斯随机矩阵。假设$\hat{}\boldsymbol{x}\in\mathbb{R}^n$ 是约束最小二乘 (1) 的一个解，且 $\boldsymbol{y}=A\boldsymbol{x}^*+\boldsymbol{e}$。如果

$$m\gtrsim m_0,$$

then

$$\|\hat{\boldsymbol{x}}-\boldsymbol{x}^*\|_2\leq\frac{3\sqrt{2}}{2(1-\rho)}\|\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-\boldsymbol{x}^*\|_2+\left(\frac{4\sqrt{2}\rho}{3(1-\rho)}+\frac{4\rho}{(1-\rho)^2}\right)\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}},\quad \textit{with}\quad \rho\gtrsim\sqrt{\frac{m_o}{m}},\tag{6}$$

对所有 $f(\boldsymbol{x}^*)\geq\eta$，以至少 $1-6\exp(-u^2)$ 的概率成立。设 $0<f(\boldsymbol{x}^*)<\eta$，且 f是绝对齐次的。记

$$\boldsymbol{x}^\eta=\frac{\eta}{f(\boldsymbol{x}^*)}\boldsymbol{x}^*,\quad m_0'=\mathcal{M}(f,\boldsymbol{x}^\eta,u),$$

If

$$m\gtrsim m_0',$$

则对于任意 $u>0$，条件

$$\|\hat{\boldsymbol{x}}-\boldsymbol{x}^*\|_2\leq\left(6\frac{\rho'}{(1-\rho')^2}+1\right)\left(\frac{\eta}{f(\boldsymbol{x}^*)}-1\right)\|\boldsymbol{x}^*\|_2+\frac{4\rho'}{(1-\rho')^2}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}},\quad \textit{with}\quad \rho'\gtrsim\sqrt{\frac{m_o}{m}},\tag{7}$$

以至少 $1-6\exp(-u^2)$ 的概率成立。

证明。 **情形 1：** $f(\boldsymbol{x}^*)\geq\eta$。由于 $\hat{}\boldsymbol{x}$ 是（1）的一个解，因此有

$$\|A\hat{\boldsymbol{x}}-\boldsymbol{y}\|_2^2\leq\|A\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-\boldsymbol{y}\|_2^2.\tag{8}$$

设

$$\boldsymbol{h}=\hat{\boldsymbol{x}}-\boldsymbol{x}^*=\boldsymbol{h}_1+\boldsymbol{h}_2$$

其中

$$\boldsymbol{h}_1=\hat{\boldsymbol{x}}-\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)\quad\text{and}\quad\boldsymbol{h}_2=\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-\boldsymbol{x}^*.$$

由式（8）可得

$$\begin{aligned}\|A\boldsymbol{h}\|_2^2 - 2\langle A^T\boldsymbol{e}, \boldsymbol{h}\rangle + \|\boldsymbol{e}\|_2^2 &= \|A\hat{\boldsymbol{x}} - A\boldsymbol{x}^* - \boldsymbol{e}\|_2^2 \\ &= \|A\hat{\boldsymbol{x}} - \boldsymbol{y}\|_2^2 \\ &\le \|A\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*) - \boldsymbol{y}\|_2^2 \\ &= \|A\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*) - A\boldsymbol{x}^* - \boldsymbol{e}\|_2^2 \\ &= \|A\boldsymbol{h}_2 - \boldsymbol{e}\|_2^2 \\ &= \|A\boldsymbol{h}_2\|_2^2 - 2\langle A^T\boldsymbol{e}, \boldsymbol{h}_2\rangle + \|\boldsymbol{e}\|_2^2.\end{aligned}$$

因此，

$$\|A\boldsymbol{h}\|_2^2 \le \|A\boldsymbol{h}_2\|_2^2 + 2\langle A^T\boldsymbol{e}, \boldsymbol{h} - \boldsymbol{h}_2\rangle = \|A\boldsymbol{h}_2\|_2^2 + 2\langle A^T\boldsymbol{e}, \boldsymbol{h}_1\rangle. \tag{9}$$

由于 $\boldsymbol{h} \in \mathcal{C}_f(\boldsymbol{x}^*)$，归一化向量 $\boldsymbol{h}/\|\boldsymbol{h}\|_2$属于集合 $T = \mathcal{C}_f(\boldsymbol{x}^*) \cap \mathcal{S}^{n-1}$。因此，应用引理 III.3，可得

$$\|A\boldsymbol{h}\|_2^2 \ge (\phi(m) - \phi(m_0))^2\|\boldsymbol{h}\|_2^2 \tag{10}$$

其概率至少为 $1 - 2\exp(-u^2)$。

接下来，由于 $f(\boldsymbol{h}_2 + \boldsymbol{x}^*) = f(\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)) \le \eta < f(\boldsymbol{x}^*)$，我们有 $\boldsymbol{h}_2 \in \mathcal{C}_f(\boldsymbol{x}^*)$。在引理 III.2 中取 $\boldsymbol{x} = \boldsymbol{h}_2/\|\boldsymbol{h}_2\|_2$，可得

$$\left|\|A\boldsymbol{h}_2\|_2 - \sqrt{m}\|\boldsymbol{h}_2\|_2\right| \le C(\omega(\mathcal{T}) + u)\|\boldsymbol{h}_2\|_2 \le C\sqrt{m_0}\|\boldsymbol{h}_2\|_2 \le \frac{\sqrt{m}}{2}\|\boldsymbol{h}_2\|_2. \tag{11}$$

其概率至少为 $1 - 2\exp(-u^2)$。由 (11) 可知

$$\|A\boldsymbol{h}_2\|_2^2 \le \frac{9m}{4}\|\boldsymbol{h}_2\|_2^2. \tag{12}$$

由于随机过程 $\langle A^T\boldsymbol{e}、\ \boldsymbol{h}_1\rangle = \langle A^T\boldsymbol{e}、\ \boldsymbol{h} - \boldsymbol{h}_2\rangle$ 中，$\boldsymbol{h}$ 和 $\boldsymbol{h}_2$ 均满足引理 III.5 中的条件，其概率至少为 $1 - 2\exp(-Cu^2)$。由 $\omega(\mathcal{T}) + u \le \sqrt{m_0}$ 可得

$$\langle A^T\boldsymbol{e}, \boldsymbol{h}_1\rangle = \langle A^T\boldsymbol{e}, \boldsymbol{h} - \boldsymbol{h}_2\rangle \le \|\boldsymbol{e}\|_2(\omega(\mathcal{T}) + u)(\|\boldsymbol{h}\|_2 + \|\boldsymbol{h}_2\|_2) \le \sqrt{m_0}\|\boldsymbol{e}\|_2(\|\boldsymbol{h}\|_2 + \|\boldsymbol{h}_2\|_2).$$

结合上述结果与 (9)、(10) 和 (12)，可得

$$\begin{aligned}(\phi(m) - \phi(m_0))^2\|\boldsymbol{h}\|_2^2 \le \|A\boldsymbol{h}\|_2^2 &\le \|A\boldsymbol{h}_2\|_2^2 + 2\langle A^T\boldsymbol{e}, \boldsymbol{h}_1\rangle \\ &\le \frac{9m}{4}\|\boldsymbol{h}_2\|_2^2 + 2\sqrt{m_0}\|\boldsymbol{e}\|_2(\|\boldsymbol{h}\|_2 + \|\boldsymbol{h}_2\|_2).\end{aligned}$$

由于引理 III.4 给出

$$1 - \rho \le 1 - \sqrt{\frac{m_0}{m}} \le 1 - \frac{\phi(m_0)}{\phi(m)},$$

我们有

$$\begin{aligned}(1-\rho)^2\|\boldsymbol{h}\|_2^2 &\le \frac{9m\|\boldsymbol{h}_2\|_2^2}{4\phi^2(m)} + \frac{2\sqrt{m_0}\|\boldsymbol{e}\|_2(\|\boldsymbol{h}\|_2 + \|\boldsymbol{h}_2\|_2)}{\phi^2(m)} \\ &\le \frac{9}{2}\|\boldsymbol{h}_2\|_2^2 + 4\rho(\|\boldsymbol{h}\|_2 + \|\boldsymbol{h}_2\|_2)\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}\end{aligned}$$

其中，最后一个不等式是利用 $\phi^2(m) \geq \frac{1.0049}{2} m$ 得出的。因此，

$$\begin{aligned}
&\left((1-\rho)\|\boldsymbol{h}\|_2 - \frac{2\rho}{1-\rho}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}\right)^2 \\
\leq &\left(\frac{3\sqrt{2}}{2}\|\boldsymbol{h}_2\|_2 + \frac{2\sqrt{2}\rho}{3}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}\right)^2 + \left(\frac{2\rho}{1-\rho}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}\right)^2 - \left(\frac{2\sqrt{2}\rho}{3}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}\right)^2 \\
\leq &\left(\frac{3\sqrt{2}}{2}\|\boldsymbol{h}_2\|_2 + \left(\frac{4\sqrt{2}\rho}{3} + \frac{2\rho}{1-\rho}\right)\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}\right)^2.
\end{aligned}$$

对上述不等式两边开平方并整理各项，得到

$$\|\boldsymbol{h}\|_2 \leq \frac{3\sqrt{2}}{2(1-\rho)}\|\boldsymbol{h}_2\|_2 + \left(\frac{4\sqrt{2}\rho}{3(1-\rho)} + \frac{2\rho}{(1-\rho)^2}\right)\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}.$$

情况 2： $0 < f(\boldsymbol{x}^*) < \eta$。由于 f是绝对齐次的，因此有 $f(\boldsymbol{x}^\eta) = \eta^{f(\boldsymbol{x}^*)f(\boldsymbol{x}^*)} = \eta$其中 $\boldsymbol{x}^\eta = {}^{\eta}_{f(\boldsymbol{x}^*)}\boldsymbol{x}^*$。令 $\tilde{\boldsymbol{e}} = A\boldsymbol{x}^* - A\boldsymbol{x}^\eta + \boldsymbol{e}$，得到

$$\boldsymbol{y} = A\boldsymbol{x}^* + \boldsymbol{e} = A\boldsymbol{x}^\eta + (A\boldsymbol{x}^* - A\boldsymbol{x}^\eta + \boldsymbol{e}) = A\boldsymbol{x}^\eta + \tilde{\boldsymbol{e}}.$$

与情形 1 类似，令 $\boldsymbol{h} = \hat{\boldsymbol{x}} - \boldsymbol{x}^* = \boldsymbol{h}_1 - \boldsymbol{h}_2$，其中 $\boldsymbol{h}_1 = \hat{\boldsymbol{x}} - \boldsymbol{x}^\eta$ 且 $\boldsymbol{h}_2 = \boldsymbol{x}^* - \boldsymbol{x}^\eta$。

由于 $\hat{\boldsymbol{x}}$是式 (1) 的一个解，因此有

$$\|A\hat{\boldsymbol{x}} - \boldsymbol{y}\|_2^2 \leq \|A\boldsymbol{x}^\eta - \boldsymbol{y}\|_2^2. \tag{13}$$

由式 (13) 可得

$$\begin{aligned}
\|A\boldsymbol{h}_1\|_2^2 - 2\langle A^T\tilde{\boldsymbol{e}}, \boldsymbol{h}_1\rangle + \|\tilde{\boldsymbol{e}}\|_2^2 &= \|A\boldsymbol{h}_1 - \tilde{\boldsymbol{e}}\|_2^2 \\
&= \|A\hat{\boldsymbol{x}} - A\boldsymbol{x}^\eta - \tilde{\boldsymbol{e}}\|_2^2 \\
&= \|A\hat{\boldsymbol{x}} - \boldsymbol{y}\|_2^2 \\
&\leq \|A\boldsymbol{x}^\eta - \boldsymbol{y}\|_2^2 = \|\tilde{\boldsymbol{e}}\|_2^2.
\end{aligned}$$

因此，

$$\|A\boldsymbol{h}_1\|_2^2 \leq 2\langle A^T\tilde{\boldsymbol{e}}, \boldsymbol{h}_1\rangle. \tag{14}$$

由于 $\boldsymbol{h}_1 \in \mathcal{C}_f(\boldsymbol{x}^\eta)$，归一化向量 $\boldsymbol{h}_1/\|\boldsymbol{h}_1\|_2$属于集合 $T' = \mathcal{C}_f(\boldsymbol{x}^\eta) \cap \mathcal{S}^{n-1}$。因此，应用引理 III.3，可得

$$\|A\boldsymbol{h}_1\|_2^2 \geq (\phi(m) - \phi(m_0'))^2\|\boldsymbol{h}_1\|_2^2. \tag{15}$$

接下来，在引理 III.5 中取 $\boldsymbol{x} = \boldsymbol{h}_1/\|\boldsymbol{h}_1\|_2$，可得

$$\langle A^T\tilde{\boldsymbol{e}}, \boldsymbol{h}_1\rangle \leq \sqrt{m_0'}\|\tilde{\boldsymbol{e}}\|_2\|\boldsymbol{h}_1\|_2. \tag{16}$$

此外，由于 $\boldsymbol{h}_2/\|\boldsymbol{h}_2\|_2$ 也属于集合 $\mathcal{T}'$，我们可以得出，概率至少为 $1 - 2\exp(-u^2)$，

$$\begin{aligned}
\|\tilde{\boldsymbol{e}}\|_2 \leq \|A\boldsymbol{h}_2\|_2 + \|\boldsymbol{e}\|_2 &\leq \left(C(\omega(\mathcal{T}') + u) + \sqrt{m}\right)\|\boldsymbol{h}_2\|_2 + \|\boldsymbol{e}\|_2 \\
&\leq \frac{3}{2}\sqrt{m}\|\boldsymbol{h}_2\|_2 + \|\boldsymbol{e}\|_2 = \frac{3}{2}\sqrt{m}\|\boldsymbol{x}^\eta - \boldsymbol{x}^*\|_2 + \|\boldsymbol{e}\|_2,
\end{aligned} \tag{17}$$

其中，在第二个不等式中，我们使用了引理 III.2；在第三个不等式中，我们使用了关于 m 的假设。结合(14)、(15)、(16)和(17)，我们得到

$$(\phi(m)-\phi(m_0'))^2\|\boldsymbol{h}_1\|_2^2 \le \|A\boldsymbol{h}_1\|_2^2 \le 2\langle A^T\tilde{\boldsymbol{e}},\boldsymbol{h}_1\rangle \le 2\sqrt{m_0'}\|\tilde{\boldsymbol{e}}\|_2\|\boldsymbol{h}_1\|_2$$
$$\le 2\sqrt{m_0'}\|\boldsymbol{h}_1\|_2\left(\frac{3}{2}\sqrt{m}\|\boldsymbol{x}^\eta-\boldsymbol{x}^*\|_2+\|\boldsymbol{e}\|_2\right).$$

最后，利用引理 III.4 和 $\phi^2(m) \ge \frac{1.0049}{2} m$，可得

$$\|\boldsymbol{h}_1\|_2 \le \frac{6\rho'}{(1-\rho')^2}\|\boldsymbol{x}^\eta-\boldsymbol{x}^*\|_2+\frac{4\rho'}{(1-\rho')^2}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}. \tag{18}$$

综上，通过 (18) 可得 (7)

$$\|\boldsymbol{h}\|_2=\|\boldsymbol{h}_1-\boldsymbol{h}_2\|_2$$
$$\le \|\boldsymbol{h}_1\|_2+\|\boldsymbol{x}^\eta-\boldsymbol{x}^*\|_2 \le \left(\frac{6\rho'}{(1-\rho')^2}+1\right)\|\boldsymbol{x}^\eta-\boldsymbol{x}^*\|_2+\frac{4\rho'}{(1-\rho')^2}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}.$$

□

注 IV.1。 定理 *IV.1* 表明，当 $f(\boldsymbol{x}^*)=\eta$ 时，(6)和(7)均由 $C\rho\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}$ 上界。~~定义~~ $\rho \ge C\sqrt{\frac{m_0}{m}}(\rho \gtrsim \sqrt{\frac{m_0}{m}})$ 中的常数$C>1$与条件 $m \ge Cm_0(m \gtrsim m_0)$~~中~~的常数一致，从而保证了分母 $1-\rho>0$。

注 IV.2。 参数 ρ 和 ρ' 均表示测量值数量与信号结构复杂度之间的比例常数。如果 f 是一个范数，则 $\rho=\rho'$。文献 *[16]* 指出，参数 ρ 和 ρ' 还决定了投影梯度算法的收敛速率。

定理 IV.1 揭示了样本复杂度（测量值的数量）与 $f(\boldsymbol{x}^*)$ 和可行集 $\mathcal{K}$ 的调节参数 η 之间一个有趣的权衡。更确切地说，如果考虑精确恢复 $\boldsymbol{y}=A\boldsymbol{x}^*$ 和 $m \ge Cm_0$，那么 $f(\boldsymbol{x}^*) \ge \eta$ 的近似误差具有如下形式

$$\|\hat{\boldsymbol{x}}-\boldsymbol{x}^*\|_2 \le \frac{3\sqrt{2}}{2\left(1-\sqrt{\frac{1}{C}}\right)}\|\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-\boldsymbol{x}^*\|_2.$$

该表达式表明，对于给定的 $\boldsymbol{x}^*$ 和 $\mathcal{K}$ （固定的结构复杂度）， m 或数据复杂度（相较于最小测量值数量 m_0）越大，近似误差就越小。项

$$\frac{3\sqrt{2}}{2(1-\rho)}\|\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-\boldsymbol{x}^*\|_2$$

(6) 中的项以及

$$\left(6\frac{\rho'}{(1-\rho')^2}+1\right)\left(\frac{\eta}{f(\boldsymbol{x}^*)}-1\right)\|\boldsymbol{x}^*\|_2$$

在（7）中，两者都源于调节参数 η 与 f（$\boldsymbol{x}^*$）之间的失配。随着 $\eta \to f$（$\boldsymbol{x}^*$），这种“失配误差”趋于零；测量值数量越多，失配误差就越小 [16]。此外，对于固定的 m， f（$\boldsymbol{x}^*$）越接近可行集 $\mathcal{K}$，近似误差就越小。请注意，本文讨论的“失配”不同于 [43], 中的“失配协方差”，后者定义为.....与

$\boldsymbol{x}$ 和残差 $\boldsymbol{y} - A\boldsymbol{x}$之间的协方差。这里我们具体量化参数 η 与真实信号结构函数 f之间的差异。

在 [5, 定理 V.1（渐近奇异性）], 中，作者研究了约束 LASSO 问题的渐近奇异性

$$\min_{\|\boldsymbol{x}\|_1 \le \eta} \|\boldsymbol{y} - A\boldsymbol{x}\|_2^2. \tag{19}$$

他们假设 $\boldsymbol{x}^*$ 是 s-稀疏的，并且 $\boldsymbol{y} = A\boldsymbol{x}^* + \sigma\boldsymbol{e}$，其中 $\boldsymbol{e} \sim \mathcal{N}(0,\ I_m)$是噪声，且 $\sigma > 0$。给定 $\eta > 0$，令$\hat{\boldsymbol{x}}$ 表示式 (19) 的解。如果 $\|\boldsymbol{x}^*\|_1 \neq \eta$ 且 $m \geq Cs\log(n/s)$，则

$$\frac{\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2^2}{\sigma^2} = \infty, \tag{20}$$

几乎必然成立。我们注意到，渐近分析 (20) 依赖于 $\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2^2$的下界，而我们在定理 IV.1 中建立了 $\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2^2$ 的非渐近误差上界。

我们从定理 IV.1 推导出若干推论，用于刻画 f 指定为给定函数时的结果。第一个结果针对标准 Lasso。

推论 IV.1。设 $\boldsymbol{x}^*$ 是 $\mathbb{R}^n$ 中的一个 s-稀疏向量。假设 $A \in \mathbb{R}^{m\times n}$ 是一个具有独立 $\boldsymbol{a}_i \sim \mathcal{N}(0, I_n)$ 行的高斯随机矩阵。令 $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ 为以下问题的一个解：

$$\min_{\|\boldsymbol{x}\|_1 \le \|\boldsymbol{x}^*\|_1} \frac{1}{2}\|\boldsymbol{y} - A\boldsymbol{x}\|_2^2 \tag{21}$$

其中 $\boldsymbol{y} = A\boldsymbol{x}^* + \boldsymbol{e}$。若 $m \gtrsim s$对数 n，则

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2 \lesssim \sqrt{\frac{s\log n}{m}}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}} \tag{22}$$

以至少 $1 - 2\exp(-s$对数 $n)$ 的概率成立。

证明。由于 $\eta = \|\boldsymbol{x}^*\|_1$，式 (6) 中的失配项为零。根据引理 10.5.3] 中的事实，有效维数 m_0 受 $s\log\ n$ 控制：任意属于 $\mathcal{C}_f(\boldsymbol{x}^*)$ 的向量都满足 $\cap\mathcal{S}^{n-1}$，其中 $\mathcal{C}_f(\boldsymbol{x}^*)$ 是 $\mathcal{D}_f(\boldsymbol{x}^*)$ 的闭锥，其

$$\mathcal{D}_f(\boldsymbol{x}^*) = \{\boldsymbol{h}:\ \|\boldsymbol{x}^* + \boldsymbol{h}\|_1 \le \|\boldsymbol{x}^*\|_1\}.$$

属于集合

$$\mathcal{T}_1^s = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}\|_1 \le \sqrt{s}, \|\boldsymbol{x}\|_2 \le 1\}. \tag{23}$$

然后，令 $u = 2\sqrt{s\log\ n}$，并假设$m \gtrsim s\log\ n$ 且 ρ 与 $\sqrt{\frac{m_0}{m}} = \sqrt{s_{\text{对数}_m}}^n$ 成正比，则可由定理 IV.1 得到估计 (22)。分母 $(1-\rho)^2$ 可视为 (0,~~1)~~中的~~某个常~~数，至此证明完毕。

□

第二个结果涉及带有 l_2 范数约束的最小二乘。

推论 IV.2。设 $\boldsymbol{x}^*$ 是 $\mathbb{R}^n$中的一个向量。假设 $A \in \mathbb{R}^{m\times n}$ 是一个具有独立$\boldsymbol{a}_i \sim \mathcal{N}(0、\ I_n)$ 行的高斯随机矩阵。设 $\hat{\boldsymbol{x}}$ 是以下问题的解

$$\min_{\|\boldsymbol{x}\|_2 \le \|\boldsymbol{x}^*\|_2} \frac{1}{2}\|\boldsymbol{y} - A\boldsymbol{x}\|_2^2, \tag{24}$$

其中 $\boldsymbol{y} = A\boldsymbol{x}^* + \boldsymbol{e}$。如果 $m \gtrsim n$，那么

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2 \lesssim \sqrt{\frac{n}{m}} \frac{\|\boldsymbol{e}\|_2}{\sqrt{m}} \tag{25}$$

以至少 $1 - 2\exp(-Cn)$ 的概率成立。

证明。与推论 IV.1 的证明类似，由于 $\eta = \|\boldsymbol{x}^*\|_2$，式 (6) 中的失配项消失。在此情况下，$m_0 \approx (\omega(\mathcal{T}) + u)^2 \approx n$，其中 $u = \sqrt{n}$。注意到当 $m \gtrsim n$ 时，我们有 $\rho = C\sqrt{\frac{n}{m}}$，且分母 $(1-\rho)^2$ 可视为 (0,1) 中的某个常数。证毕。 □

我们注意到，推论 IV.1 已在 [44] 和 [39, 定理 10.6.1], 中得到证明，而推论 IV.2 的结果则给出了与 [1] 中线性估计所得到的相同上界。

从推论 IV.1 可以推知，我们的结果将经典压缩感知结果的类似结论作为特例纳入其中，即 $f(\boldsymbol{x}^*) = \|\boldsymbol{x}^*\|_1 = \eta$。对于非凸约束，也可以得到与推论 IV.1 一致的结论

$$\mathcal{K} = \{\boldsymbol{x} \in \mathbb{R}^n : \ \|\boldsymbol{x}\|_p \le \|\boldsymbol{x}^*\|_p = \eta\}, \quad 0 \le p < 1$$

基于高斯宽度的以下性质： $\omega(\mathcal{T}) = \omega(\text{conv}(\mathcal{T}))$。我们断言，定理 IV.1 有多种证明方法。具体而言，可以利用引理 III.1 推导出另一种证明，从而得到一个结果，其特征是一个由 δ 表示的乘法因子，如下所示。

定理 IV.2. 设 $\boldsymbol{x}^*$ 是 $\mathbb{R}^n$ 中的任意向量。对于任意 $u > 0$，令 $m_0 = \mathcal{M}(f, \boldsymbol{x}^*, u)$，其定义见定义 *III.3*。设 $A \in \mathbb{R}^{m\times n}$ 为一个高斯随机矩阵，其各行是独立的 $\boldsymbol{a}_i \sim \mathcal{N}(0, I_n)$。假设 $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ 是约束最小二乘 (1) 的一个解，且 $\boldsymbol{y} = A\boldsymbol{x}^* + \boldsymbol{e}$。若

$$m \ge \frac{1}{\delta^2} m_0, \quad \textit{with} \quad \delta \in (0,1)$$

then

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2 \le \frac{\sqrt{2}(1+\delta)}{(1-\rho)} \|\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*) - \boldsymbol{x}^*\|_2 + \left(\frac{4\sqrt{2}\rho}{3(1-\rho)} + \frac{4\rho}{(1-\rho)^2}\right) \frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}, \quad \textit{with} \quad \rho = \frac{1}{\delta}\sqrt{\frac{m_o}{m}},$$

对于所有 $f(\boldsymbol{x}^*) \ge \eta$ 均成立，且其概率至少为 $1 - 6\exp(-u^2)$。假设 $0 < f(\boldsymbol{x}^*) < \eta$，并且 f 是绝对齐次的。记

$$\boldsymbol{x}^\eta = \frac{\eta}{f(\boldsymbol{x}^*)} \boldsymbol{x}^*, \quad m_0' = \mathcal{M}(f, \boldsymbol{x}^\eta, u),$$

If

$$m \ge \frac{1}{\delta^2} m_0', \quad \textit{with} \quad \delta \in (0,1),$$

那么，对于任意 $u > 0$，条件

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2 \le \left(4\frac{\rho'(1+\delta)}{(1-\rho')^2} + 1\right)\left(\frac{\eta}{f(\boldsymbol{x}^*)} - 1\right)\|\boldsymbol{x}^*\|_2 + \frac{4\rho'}{(1-\rho')^2}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}, \quad \textit{with} \quad \rho = \frac{1}{\delta}\sqrt{\frac{m_o}{m}},$$

以至少 $1 - 6\exp(-u^2)$ 的概率成立。

定理 IV.2 的证明见附录。下面给出约束最小绝对偏差 (2) 的解的估计。

定理 IV.3. 设 $\boldsymbol{x}^*$ 是 $\mathbb{R}^n$ 中任意一个向量。设 $m_1 = \mathcal{M}(f, \boldsymbol{x}^*)$，其定义见定义 *III.4*。设$\gamma \geq 4$, $\beta > (\frac{5}{4}\gamma)^2$。假设 $A \in \mathbb{R}^{m\times n}$ 是一个高斯随机矩阵，其各行相互独立，且服从 $\boldsymbol{a}_i \sim \mathcal{N}(0, I_n)$ 分布。设$\hat{}\boldsymbol{x}$ 是满足 $\boldsymbol{y} = A\boldsymbol{x}^* + \boldsymbol{e}$ 的约束最小绝对偏差(2) 的一个解。如果

$$m > \beta m_1, \tag{26}$$

那么对于 $0 < u < \sqrt{\frac{2}{\pi}} - \rho$,

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2 \leq \frac{\sqrt{\frac{2}{\pi}} + \rho + u}{\sqrt{\frac{2}{\pi}} - \rho - u}\|\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*) - \boldsymbol{x}^*\|_2 + \frac{2}{\sqrt{\frac{2}{\pi}} - \rho - u}\frac{\|\boldsymbol{e}\|_1}{m}, \quad \textit{with} \quad \rho = \gamma\sqrt{\frac{m_1}{m}}, \tag{27}$$

以至少 $1 - 6\exp(-\frac{mu^2}{2})$ 的概率，对所有 $f(\boldsymbol{x}^*) \geq \eta$均成立。假设 $0 < f(\boldsymbol{x}^*) < \eta$，且 f 是绝对齐次的。记

$$\boldsymbol{x}^\eta = \frac{\eta}{f(\boldsymbol{x}^*)}\boldsymbol{x}^*, \quad m_1' = \mathcal{M}(f, \boldsymbol{x}^\eta).$$

If

$$m > \beta m_1', \tag{28}$$

那么对于任意 $0 < u < \sqrt{\frac{2}{\pi}} - \rho'$,

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2 \leq \frac{3\sqrt{\frac{2}{\pi}} + \rho' + u}{\sqrt{\frac{2}{\pi}} - \rho' - u}\left(\frac{\eta}{f(\boldsymbol{x}^*)} - 1\right)\|\boldsymbol{x}^*\|_2 + \frac{2}{\sqrt{\frac{2}{\pi}} - \rho' - u}\frac{\|\boldsymbol{e}\|_1}{m}, \quad \textit{with} \quad \rho' = \gamma\sqrt{\frac{m_1'}{m}}, \tag{29}$$

以至少 $1 - 6\exp(-\frac{mu^2}{2})$ 的概率成立。 ——

定理 IV.3 的证明见附录。定理 IV.3 还揭示了“失配误差”与数据复杂度之间的权衡。我们关注 $f(\boldsymbol{x}^*) < \eta$ 的情形。令 ϵ 为最优解所需的相对精度，则有

$$\frac{\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2}{\|\boldsymbol{x}^*\|_2} \leq \epsilon = \frac{3\sqrt{\frac{2}{\pi}} + \rho' + u}{\sqrt{\frac{2}{\pi}} - \rho' - u}\left(\frac{\eta}{f(\boldsymbol{x}^*)} - 1\right) + \frac{2}{\sqrt{\frac{2}{\pi}} - \rho' - u}\frac{\|\boldsymbol{e}\|_1}{m\|\boldsymbol{x}^*\|_2}.$$

可以观察到，测量值的数量越多，速率 ρ'的值就越小，进而导致更小的失配误差

$$\frac{3\sqrt{\frac{2}{\pi}} + \rho' + u}{\sqrt{\frac{2}{\pi}} - \rho' - u}\left(\frac{\eta}{f(\boldsymbol{x}^*)} - 1\right).$$

另一方面,当噪声水平 $\|\boldsymbol{e}\|_1$ 相较于 $\|\boldsymbol{x}^*\|_2$足够小时，不等式

$$\frac{3\sqrt{\frac{2}{\pi}} + \rho' + u}{\sqrt{\frac{2}{\pi}} - \rho' - u}\left(\frac{\eta}{f(\boldsymbol{x}^*)} - 1\right) \leq \epsilon$$

成立。简单变换可得

$$m \geq m_0\frac{(\text{mismatch} + \epsilon)^2}{(C_2\epsilon - C_1 \cdot \text{mismatch})^2}, \tag{30}$$

其中 $C_1 = 3\sqrt{\frac{2}{\pi}} + u$, $C_2 = \sqrt{\frac{2}{\pi}} - u$，我们用"失配"表示失配项 $f^{\eta}(\boldsymbol{x}^*) - 1$。表达式 (30) 表明，如果项 $f^{\eta}_{\boldsymbol{x}^*}() - 1$ 趋于零，则所需的数据更少（所需的 m 更小）。接下来，我们考虑 $f(\boldsymbol{x}) = \|\boldsymbol{x}\|_1$ 的稀疏情形。

推论 IV.3。令 $\boldsymbol{x}^*$ 为 $\mathbb{R}^n$ 中任意一个 s-稀疏向量。令 $\gamma \geq 4$, $\beta > (\frac{5}{4}\gamma)^2$。假设 $A \in \mathbb{R}^{m\times n}$ 是一个具有独立 $\boldsymbol{a}_i \sim \mathcal{N}(0, I_n)$ 行的高斯随机矩阵。令 $\hat{\boldsymbol{x}}$ 为约束最小绝对偏差的一个解：

$$\min_{\|\boldsymbol{x}\|_1 \leq \|\boldsymbol{x}^*\|_1} \|\boldsymbol{y} - A\boldsymbol{x}\|_1 \tag{31}$$

其中 $\boldsymbol{y} = A\boldsymbol{x}^* + \boldsymbol{e}$。若

$$m > \beta s \log n,$$

则对于 $0 < u < \sqrt{\frac{2}{\pi}} - \gamma\sqrt{\frac{s \log n}{m}}$，

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2 \leq \frac{2}{\sqrt{\frac{2}{\pi}} - u - \gamma\sqrt{\frac{s\log n}{m}}} \frac{\|\boldsymbol{e}\|_1}{m} \tag{32}$$

以至少 $1 - 2\exp(-\frac{mu^2}{2})$ 的概率成立。

证明。由于 $\eta = \|\boldsymbol{x}^*\|_1$，"失配误差"项为零。由于约束集与推论 IV.1 中的相同，有效维数 m_1 也受 $s\log n$ 和 $\rho = \gamma\sqrt{\frac{m_1}{m}} \approx \sqrt{s\log_m n}$ 控制。于是，我们得到结论 (32)。

□

如果观测值受到对抗性损坏，即，

$$\boldsymbol{y} = A\boldsymbol{x}^* + \boldsymbol{e}_1 + \boldsymbol{e}_2$$

其中 $\|\boldsymbol{x}^*\|_0 \leq s$，且 $\|\boldsymbol{e}_1\|_0 \leq \beta m$，那么 [7] 中已证明，约束最小绝对偏差 (31) 对任意比例的损坏 β 小于 $\beta_0 \approx 0.239$均具有鲁棒性。我们注意到，$\|\boldsymbol{x}^*\|_1 \leq \eta$ 的情形仅在 [7] 中进行了研究，而 $\|\boldsymbol{x}^*\|_1 > \eta$ 的情形尚未解决。设 $\hat{\boldsymbol{x}}$ 为以下约束最小绝对偏差问题的一个解

$$\min_{\|\boldsymbol{x}\|_1 \leq \eta} \|\boldsymbol{y} - A\boldsymbol{x}\|_1.$$

结合 [7,定理 1.1] 证明中的方法与定理 IV.3证明中的技术，我们得到，对于 $\|\boldsymbol{x}^*\|_1 > \eta$，以下结论

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2 \lesssim \frac{1}{\epsilon - \frac{1}{\alpha}} \left(\frac{1}{m}\|\boldsymbol{e}_2\|_1 + \left(\sqrt{\frac{1}{2\pi}} + \frac{\epsilon}{2}\right) \|\mathcal{P}_{\mathcal{K}}(\boldsymbol{x}^*) - \boldsymbol{x}^*\|_2 \right) + \frac{\|\boldsymbol{x}^*\|_1 - \eta}{\alpha\sqrt{s}}, \tag{33}$$

只要满足以下条件，就以高概率成立

$$m \gtrsim \frac{\alpha^2}{\epsilon^2} s \log\left(\frac{en}{\alpha^2 \epsilon s}\right),$$

其中 $\epsilon > 0$、$\beta < \beta_0 - \epsilon$ 具有 $\beta_0 \approx 0.239$ 和 $\alpha \geq \frac{2}{\epsilon}$。该估计 (33) 可作为对 [7, 定理 1.1] 的补充。

B. 相位恢复

定理 IV.4。设 $\boldsymbol{x}^*$ 是 $\mathbb{R}^n$ 中的任意向量。对于任意 $u>0$，令 $m_0=\mathcal{M}(f、\boldsymbol{x}^*、u)$，其定义见定义 *III.3*。假设 $A\in\mathbb{R}^{m\times n}$ 是一个高斯随机矩阵，其行相互独立，服从 $\boldsymbol{a}_i\sim\mathcal{N}(0、I_n)$。令$\hat{}\boldsymbol{x}$ 为满足 (3) 且具有 $\boldsymbol{y}=|A\boldsymbol{x}^*|+\boldsymbol{e}$ 的一个解。如果

$$m\gtrsim m_0,$$

then

$$\min\{\|\hat{\boldsymbol{x}}-\boldsymbol{x}^*\|_2,\|\hat{\boldsymbol{x}}+\boldsymbol{x}^*\|_2\}\le\left(\frac{4}{v_0}\rho+\frac{4}{v_0}+1\right)\|\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-\boldsymbol{x}^*\|_2+\frac{4}{v_0}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}},\quad \textit{with}\quad \rho\gtrsim\sqrt{\frac{m_0}{m}},\tag{34}$$

以至少为 $1-2\exp(-u^2)-2\exp\left(\frac{v_0^2}{8}m\right)$ 的概率对所有 $f(\boldsymbol{x}^*)\ge\eta$成立。设 f是绝对齐次的，并满足 $0<f(\boldsymbol{x}^*)<\eta$。记

$$m_0'=\mathcal{M}(f,\boldsymbol{x}^\eta,u),\quad v_0=\frac{1}{32e}\sqrt{\frac{\pi}{2}}\left(1-\frac{1}{4\sqrt{\pi}}\right)\approx 0.0124.$$

If

$$m\gtrsim m_0',$$

则对于任意 $u>0$,

$$\min\{\|\hat{\boldsymbol{x}}-\boldsymbol{x}^*\|_2,\|\hat{\boldsymbol{x}}+\boldsymbol{x}^*\|_2\}\le\left(\frac{24}{v_0^2}\rho'+\frac{3}{v_0}+1\right)\left(\frac{\eta}{f(\boldsymbol{x}^*)}-1\right)\|\boldsymbol{x}^*\|_2+\left(\frac{16}{v_0^2}\rho'+\frac{2}{v_0}\right)\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}},\quad \textit{with}\quad \rho'\gtrsim\sqrt{\frac{m_0'}{m}},\tag{35}$$

以至少为 $1-2\exp(-u^2)-2\exp\left(\frac{v_0^2}{8}m\right)$. 的概率成立

证明。 **情形 1：** $f(\boldsymbol{x}^*)\ge\eta$。 记指标集

$$S_1=\{j:\langle\boldsymbol{a}_j,\hat{\boldsymbol{x}}\rangle\langle\boldsymbol{a}_j,\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)\rangle>0\},\quad S_2=\{j:\langle\boldsymbol{a}_j,\hat{\boldsymbol{x}}\rangle\langle\boldsymbol{a}_j,\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)\rangle<0\}.$$

不失一般性地，我们假设 $\sharp(S_1)=\beta m\ge m/2$ （否则，我们可以假设 $\sharp(S_2)\ge m/2$）。令 $\boldsymbol{h}^-=\hat{}\boldsymbol{x}-\boldsymbol{x}^*$ 和 $\boldsymbol{h}^+=\hat{}\boldsymbol{x}+\boldsymbol{x}^*$。类似于定理 IV.1 的证明，记

$$\boldsymbol{h}^-=\boldsymbol{h}_1+\boldsymbol{h}_2,$$

其中

$$\boldsymbol{h}_1=\hat{\boldsymbol{x}}-\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)\quad\text{and}\quad\boldsymbol{h}_2=\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-\boldsymbol{x}^*.$$

对于 S_1，我们有

$$\begin{aligned}\|A_{S_1}\boldsymbol{h}_1\|_2 &= \|A_{S_1}\hat{\boldsymbol{x}} - A_{S_1}\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)\|_2\\ &= \||A_{S_1}\hat{\boldsymbol{x}}| - |A_{S_1}\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)|\|_2\\ &\le \||A\hat{\boldsymbol{x}}| - |A\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)|\|_2\\ &\le \||A\hat{\boldsymbol{x}}| - \boldsymbol{y}\|_2 + \||A\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)| - \boldsymbol{y}\|_2\\ &\le 2\||A\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)| - \boldsymbol{y}\|_2\\ &= 2\||A\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)| - |A\boldsymbol{x}^*| - \boldsymbol{e}\|_2\\ &\le 2\||A\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)| - |A\boldsymbol{x}^*|\|_2 + 2\|\boldsymbol{e}\|_2\\ &\le 2\|A\boldsymbol{h}_2\|_2 + 2\|\boldsymbol{e}\|_2, \end{aligned} \tag{36}$$

其中，第五个不等式基于$\hat{\boldsymbol{x}}$ 是式 (3) 的解这一事实成立。

由于 $f(\boldsymbol{h}_2+\boldsymbol{x}^*) = f(\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)) \le \eta \le f(\boldsymbol{x}^*)$，归一化向量 $\boldsymbol{h}_2/\|\boldsymbol{h}_2\|_2$属于集合 $T = \mathcal{C}_f(\boldsymbol{x}^*)\cap\mathcal{S}^{n-1}$，因此引理 III.2 推出

$$\|A\boldsymbol{h}_2\|_2 \le (C(\omega(\mathcal{T}) + u) + \sqrt{m})\|\boldsymbol{h}_2\|_2 \le C(\sqrt{m_0} + \sqrt{m})\|\boldsymbol{h}_2\|_2 \tag{37}$$

以至少 $1 - 2\exp(-u^2)$ 的概率成立。

此外，引理 III.6 推出

$$\|A_{S_1}\boldsymbol{h}_1\|_2 \ge \frac{v_0}{2}\sqrt{m}\|\boldsymbol{h}_1\|_2 \tag{38}$$

以至少 $1 - 2\exp\left(-\frac{v_0^2}{8}m\right)$ 的概率成立。

因此，由（36）、（37）和（38），我们有

$$\begin{aligned}\|\boldsymbol{h}^-\|_2 = \|\boldsymbol{h}_1 + \boldsymbol{h}_2\|_2 &\le \|\boldsymbol{h}_1\|_2 + \|\boldsymbol{h}_2\|_2\\ &\le \frac{2}{v_0\sqrt{m}}\|A_{S_1}\boldsymbol{h}_1\|_2 + \|\boldsymbol{h}_2\|_2\\ &\le \frac{4}{v_0\sqrt{m}}(\|A\boldsymbol{h}_2\|_2 + \|\boldsymbol{e}\|_2) + \|\boldsymbol{h}_2\|_2\\ &\le \frac{4}{v_0\sqrt{m}}\left(C\sqrt{m_0} + \sqrt{m}\right)\|\boldsymbol{h}_2\|_2 + \frac{4}{v_0\sqrt{m}}\|\boldsymbol{e}\|_2 + \|\boldsymbol{h}_2\|_2\\ &\le \left(\frac{4}{v_0}\rho + \frac{4}{v_0} + 1\right)\|\boldsymbol{h}_2\|_2 + \frac{4}{v_0}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}\end{aligned}$$

以至少为 $1 - 2\exp(-u^2) - 2\exp\left(\frac{v_0^2}{8}m\right)$的概率成立。对于 $\sharp(S_2) \ge m/2$ 的情况，使用类似的方法，我们可以得出 $\boldsymbol{h}^+$ 具有相同的上界。

情况 2： $0 < f(\boldsymbol{x}^*) < \eta$。我们记 $\boldsymbol{x}^\eta = \frac{\eta}{f(\boldsymbol{x}^*)}\boldsymbol{x}^*$和 $\tilde{\boldsymbol{e}} = |A\boldsymbol{x}^*| - |A\boldsymbol{x}^\eta| + \boldsymbol{e}$。由于 f 是绝对齐次的，我们有

$$f(\boldsymbol{x}^\eta) = \eta\frac{f(\boldsymbol{x}^*)}{f(\boldsymbol{x}^*)} = \eta.$$

此外，

$$\boldsymbol{y} = |A\boldsymbol{x}^*| + \boldsymbol{e} = |A\boldsymbol{x}^\eta| + (|A\boldsymbol{x}^*| - |A\boldsymbol{x}^\eta| + \boldsymbol{e}) = |A\boldsymbol{x}^\eta| + \tilde{\boldsymbol{e}},$$

其中， $\tilde{\boldsymbol{e}}$ 可视为测量值 $|A\boldsymbol{x}^{\eta}|$ 上的加性噪声。

由于 $\hat{\boldsymbol{x}}$ 是 (3) 的解，我们有

$$\||A\hat{\boldsymbol{x}}| - \boldsymbol{y}\|_2^2 \leq \||A\boldsymbol{x}^{\eta}| - \boldsymbol{y}\|_2^2. \tag{39}$$

不同于情形 1，我们设

$$\begin{aligned}
S_1 &= \{j : \operatorname{sign}(\langle \boldsymbol{a}_j, \hat{\boldsymbol{x}}\rangle) = 1, \operatorname{sign}(\langle \boldsymbol{a}_j, \boldsymbol{x}^{\eta}\rangle) = 1\},\\
S_2 &= \{j : \operatorname{sign}(\langle \boldsymbol{a}_j, \hat{\boldsymbol{x}}\rangle) = -1, \operatorname{sign}(\langle \boldsymbol{a}_j, \boldsymbol{x}^{\eta}\rangle) = -1\},\\
S_3 &= \{j : \operatorname{sign}(\langle \boldsymbol{a}_j, \hat{\boldsymbol{x}}\rangle) = 1, \operatorname{sign}(\langle \boldsymbol{a}_j, \boldsymbol{x}^{\eta}\rangle) = -1\},\\
S_4 &= \{j : \operatorname{sign}(\langle \boldsymbol{a}_j, \hat{\boldsymbol{x}}\rangle) = -1, \operatorname{sign}(\langle \boldsymbol{a}_j, \boldsymbol{x}^{\eta}\rangle) = 1\}.
\end{aligned}$$

不失一般性，假设 $|(S_1 \cup S_2)| \geq m/2$。我们还记

$$\boldsymbol{h}^- = \hat{\boldsymbol{x}} - \boldsymbol{x}^* \quad \text{and} \quad \boldsymbol{h}^+ = \hat{\boldsymbol{x}} + \boldsymbol{x}^*.$$

此外，我们设

$$\boldsymbol{h}^- = \boldsymbol{h}_1 + \boldsymbol{h}_2,$$

其中

$$\boldsymbol{h}_1 = \hat{\boldsymbol{x}} - \boldsymbol{x}^{\eta} \quad \text{and} \quad \boldsymbol{h}_2 = \boldsymbol{x}^{\eta} - \boldsymbol{x}^*.$$

Then

$$\||A\hat{\boldsymbol{x}}| - \boldsymbol{y}\|_2^2 = \||A\hat{\boldsymbol{x}}| - |A\boldsymbol{x}^{\eta}| - \tilde{\boldsymbol{e}}\|_2^2 \geq \|A_{S_1}\boldsymbol{h}_1 - \tilde{\boldsymbol{e}}_{S_1}\|_2^2 + \|A_{S_2}\boldsymbol{h}_1 + \tilde{\boldsymbol{e}}_{S_2}\|_2^2.$$

结合 (39)，可得

$$\|A_{S_1}\boldsymbol{h}_1 - \tilde{\boldsymbol{e}}_{S_1}\|_2^2 + \|A_{S_2}\boldsymbol{h}_1 + \tilde{\boldsymbol{e}}_{S_2}\|_2^2 \leq \|\tilde{\boldsymbol{e}}\|_2^2. \tag{40}$$

展开 (40) 中的平方项并重新排列各项，可得

$$\|A_{S_{12}}\boldsymbol{h}_1\|_2^2 \leq 2\langle \boldsymbol{h}_1, A_{S_1}^T\tilde{\boldsymbol{e}}_{S_1} - A_{S_2}^T\tilde{\boldsymbol{e}}_{S_2}\rangle + \|\tilde{\boldsymbol{e}}_{S_{12}^c}\|_2^2. \tag{41}$$

此外，由引理 III.6 可知

$$\|A_{S_{12}}\boldsymbol{h}_1\|_2^2 \geq \frac{v_0^2}{4}m\|\boldsymbol{h}_1\|_2^2 \tag{42}$$

以至少 $1 - 2\exp\left(-\frac{v_0^2}{8}m\right)$ 的概率成立。

接下来，我们估计 (41) 中的内积项

$$\left\langle \boldsymbol{h}_1, A_{S_1}^T\tilde{\boldsymbol{e}}_{S_1} - A_{S_2}^T\tilde{\boldsymbol{e}}_{S_2}\right\rangle.$$

由于 $f(\boldsymbol{h}_1 + \boldsymbol{x}^{\eta}) = f(\hat{\boldsymbol{x}}) \leq \eta = f(\boldsymbol{x}^{\eta})$，因此 $\boldsymbol{h}_1/\|\boldsymbol{h}_1\|_2$ 的归一化向量属于集合 $\mathcal{T}' = \mathcal{C}_f(\boldsymbol{x}^{\eta}) \cap \mathcal{S}^{n-1}$。利用引理 III.5，可得

$$\langle \boldsymbol{h}_1, A_{S_1}^T\tilde{\boldsymbol{e}}_{S_1} - A_{S_2}^T\tilde{\boldsymbol{e}}_{S_2}\rangle \leq 2(\omega(\mathcal{T}') + u)\|\tilde{\boldsymbol{e}}\|_2\|\boldsymbol{h}_1\|_2 \leq 2\sqrt{m_0'}\|\tilde{\boldsymbol{e}}\|_2\|\boldsymbol{h}_1\|_2, \tag{43}$$

以至少 $1 - 2\exp(-Cu^2)$ 的概率成立。

将 (42)、(43) 代入方程 (41)，可得

$$\begin{aligned}\frac{v_0^2 m}{4}\|\boldsymbol{h}_1\|_2^2 &\le \|A_{S_{12}}\boldsymbol{h}_1\|_2^2\\ &\le 2\langle \boldsymbol{h}_1, A_{S_1}^T\tilde{\boldsymbol{e}}_{S_1} - A_{S_2}^T\tilde{\boldsymbol{e}}_{S_2}\rangle + \|\tilde{\boldsymbol{e}}_{S_{12}^c}\|_2^2\\ &\le 4\sqrt{m_0'}\|\tilde{\boldsymbol{e}}\|_2\|\boldsymbol{h}_1\|_2 + \|\tilde{\boldsymbol{e}}_{S_{12}^c}\|_2^2,\end{aligned}$$

对不等式两边配方后，可得

$$\left(\frac{v_0}{2}\sqrt{m}\|\boldsymbol{h}_1\|_2 - 4\sqrt{\frac{m_0'}{m}}\frac{\|\tilde{\boldsymbol{e}}\|_2}{v_0}\right)^2 \le \|\tilde{\boldsymbol{e}}\|_2^2 + \left(4\sqrt{\frac{m_0'}{m}}\frac{\|\tilde{\boldsymbol{e}}\|_2}{v_0}\right)^2 \le \left(\left(v_0 + 4\sqrt{\frac{m_0'}{m}}\right)\frac{\|\tilde{\boldsymbol{e}}\|_2}{v_0}\right)^2, \tag{44}$$

其中，最后一个不 等式由 $a^2+b^2 \le (a+b)^2$ 得出。对 (44) 两边开平方，得到 s

$$\|\boldsymbol{h}_1\|_2 \le \frac{2}{v_0}\left(\frac{8}{v_0}\sqrt{\frac{m_0'}{m}} + 1\right)\frac{\|\tilde{\boldsymbol{e}}\|_2}{\sqrt{m}}. \tag{45}$$

下面，我们通过以下方式估计 $\|\tilde{\boldsymbol{e}}\|_2$

$$\begin{aligned}\|\tilde{\boldsymbol{e}}\|_2 &= \||A\boldsymbol{x}^*| - |A\boldsymbol{x}^\eta| + \boldsymbol{e}\|_2\\ &\le \||A\boldsymbol{x}^*| - |A\boldsymbol{x}^\eta|\|_2 + \|\boldsymbol{e}\|_2\\ &\le \|A(\boldsymbol{x}^* - \boldsymbol{x}^\eta)\|_2 + \|\boldsymbol{e}\|_2\end{aligned} \tag{46}$$

根据关于 m 的假设，并在引理 III.2 中取 $\boldsymbol{x} = \frac{\boldsymbol{x}^*-\boldsymbol{x}^\eta}{\|\boldsymbol{x}^*-\boldsymbol{x}^\eta\|_2}$，可得

$$\begin{aligned}\|A(\boldsymbol{x}^* - \boldsymbol{x}^\eta)\|_2 &\le \left(C(\omega(\mathcal{T}) + u) + \sqrt{m}\right)\|(\boldsymbol{x}^* - \boldsymbol{x}^\eta)\|_2\\ &\le \left(C\sqrt{m_0'} + \sqrt{m}\right)\|(\boldsymbol{x}^* - \boldsymbol{x}^\eta)\|_2 \le \frac{3\sqrt{m}}{2}\|(\boldsymbol{x}^* - \boldsymbol{x}^\eta)\|_2.\end{aligned} \tag{47}$$

因此，由 (46) 和 (47)，我们得到

$$\|\tilde{\boldsymbol{e}}\|_2 \le \frac{3}{2}\sqrt{m}\|\boldsymbol{x}^* - \boldsymbol{x}^\eta\|_2 + \|\boldsymbol{e}\|_2. \tag{48}$$

将 (48) 代入 (45)，可得

$$\begin{aligned}\|\boldsymbol{h}_1\|_2 &\le \left(\frac{24}{v_0^2}\sqrt{\frac{m_0'}{m}} + \frac{3}{v_0}\right)\|\boldsymbol{x}^* - \boldsymbol{x}^\eta\|_2 + \left(\frac{16}{v_0^2}\sqrt{\frac{m_0'}{m}} + \frac{2}{v_0}\right)\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}\\ &= \left(\frac{24}{v_0^2}\sqrt{\frac{m_0'}{m}} + \frac{3}{v_0}\right)\left(\frac{\eta}{f(\boldsymbol{x}^*)} - 1\right)\|\boldsymbol{x}^*\|_2 + \left(\frac{16}{v_0^2}\sqrt{\frac{m_0'}{m}} + \frac{2}{v_0}\right)\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}.\end{aligned} \tag{49}$$

由 (49) 和三角不等式即可完成证明

$$\begin{aligned}\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2 = \|\boldsymbol{h}^-\|_2 &\le \|\boldsymbol{h}_1\|_2 + \|\boldsymbol{h}_2\|_2\\ &= \|\hat{\boldsymbol{x}} - \boldsymbol{x}^\eta\|_2 + \|\boldsymbol{x}^\eta - \boldsymbol{x}^*\|_2\\ &\le \left(\frac{24}{v_0^2}\sqrt{\frac{m_0'}{m}} + \frac{3}{v_0} + 1\right)\left(\frac{\eta}{f(\boldsymbol{x}^*)} - 1\right)\|\boldsymbol{x}^*\|_2 + \left(\frac{16}{v_0^2}\sqrt{\frac{m_0'}{m}} + \frac{2}{v_0}\right)\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}.\end{aligned}$$

对于 $|(S_3 \cup S_4)| \ge m/2$ 的情况，我们可以用与上述类似的方法得到 $\|\boldsymbol{h}^+\|_2$ 的相同界。 □

当参数取最优值时，我们得到关于稀疏相位恢复问题的以下两个推论。

推论 IV.4. 设 $\boldsymbol{x}^*$ 是 $\mathbb{R}^n$ 中任意一个 s-稀疏向量。假设 $A \in \mathbb{R}^{m\times n}$ 是一个高斯随机矩阵，其各行独立，服从 $\boldsymbol{a}_i \sim \mathcal{N}(0,\ I_n)$ 分布。设 $\hat{}\boldsymbol{x} \in \mathbb{R}^n$ 是以下问题的一个解

$$\min_{\|\boldsymbol{x}\|_1 \le \|\boldsymbol{x}^*\|_1} \frac{1}{2}\left\|\boldsymbol{y} - |A\boldsymbol{x}|\right\|_2^2 \tag{50}$$

其中 $\boldsymbol{y} = |A\boldsymbol{x}^*| + \boldsymbol{e}$。如果

$$m \gtrsim s\log n,$$

then

$$\min\{\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2, \|\hat{\boldsymbol{x}} + \boldsymbol{x}^*\|_2\} \lesssim \frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}$$

以至少 $1 - 3\exp(-Cm)$ 的概率成立。

证明。 由于 $\eta = \|\boldsymbol{x}^*\|_1$，式 (34) 和 (35) 中的"失配误差"项消失。与推论 IV.1 的证明类似，有效维数 m_0 受 $s\log\ n$ 控制。注意，当 $m \gtrsim s\log\ n$ 时， ρ' 的值位于区间 (0,1) 内，因此式 (34) 中的 $\frac{4}{v_0}$ 项和式 (35) 中的 $\frac{16}{v_0^2}\rho' + \frac{2}{v_0}$ 项都可以视为某个常数。于是，根据定理 IV.4 可得

$$\min\{\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2, \|\hat{\boldsymbol{x}} + \boldsymbol{x}^*\|_2\} \le \frac{C}{v_0}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}.$$

□

推论 IV.4 与 与 [32, 定理 I.5], 一致，并且在 [32, 注释 I.4] 中证明了界限 $\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}$ 是尖锐的 .

推论 IV.5. 设 $\boldsymbol{x}^*$ 是 $\mathbb{R}^n$ 中任意一个 s-稀疏向量。假设 $A \in \mathbb{R}^{m\times n}$ 是一个具有独立 $\boldsymbol{a}_i \sim \mathcal{N}(0, I_n)$ 行的高斯随机矩阵。设 $\hat{}\boldsymbol{x} \in \mathbb{R}^n$ 是以下问题的一个解

$$\min_{\|\boldsymbol{x}\|_0 \le s} \frac{1}{2}\left\|\boldsymbol{y} - |A\boldsymbol{x}|\right\|_2^2 \tag{51}$$

其中 $\boldsymbol{y} = |A\boldsymbol{x}^*| + \boldsymbol{e}$。如果

$$m \gtrsim s\log n,$$

then

$$\min\{\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|_2, \|\hat{\boldsymbol{x}} + \boldsymbol{x}^*\|_2\} \lesssim \frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}$$

以至少 $1 - 3\exp(-Cm)$ 的概率成立。

证明。 结合以下事实 $\omega(\mathcal{T}_1^s) \le 2\omega(\mathcal{T}_0^s) \le C\sqrt{s\log\ n}$，其中 $\mathcal{T}_0^s = \{\boldsymbol{x}:\ \|\boldsymbol{x}\|_0 \le s, \|\boldsymbol{x}\|_2 \le 1\}$，以及推论 IV.4 的证明，即可直接得到此证明。 □

注释 IV.3. 对于稀疏相位恢复问题(51)，迭代硬阈值*[35]*,截断幅度流 *[30]*,*CoPRAM [36]* 等迭代算法通常包含两个阶段：初始化和细化。初始化阶段采用谱初始化，需要 $\mathcal{O}(s^2\log n)$ 个高斯样本，以获得足够准确的估计。在细化阶段，各种算法会进一步优化该估计，其中大多数算法仅需 $\mathcal{O}(s\log\ n)$ 个样本即可实现线性收敛。因此，总样本复杂度主要由初始化阶段决定

阶段。模型在使用 $\mathcal{O}(s\log\ n)$ 个测量值时所提供的恢复保证，与算法实际样本复杂度（因初始化而增加）之间的差距，是后续工作需要解决的一项关键挑战。

比较定理 IV.4 和推论 IV.4 可知，存在额外项

$$\left(\frac{4}{v_0}\rho+\frac{4}{v_0}+1\right)\|\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-\boldsymbol{x}^*\|_2,\quad \text{for}\quad f(\boldsymbol{x}^*)\geq\eta,$$

and

$$\left(\frac{24}{v_0^2}\rho'+\frac{3}{v_0}+1\right)\left(\frac{\eta}{f(\boldsymbol{x}^*)}-1\right)\|\boldsymbol{x}^*\|_2,\quad \text{for}\quad 0<f(\boldsymbol{x}^*)<\eta.$$

这些额外项与线性情形中的项一致，源于调节参数η 与 $f(\boldsymbol{x}^*)$ 之间的失配。定理 IV.4 表明，随着数据复杂度 m 增大，速率 ρ 或 ρ' 会降低，从而使失配误差减小。定理 IV.4 还展示了“失配误差”与数据复杂度之间的权衡。为此，考虑 $f(\boldsymbol{x}^*)\geq\eta$ 的情形，我们用 ϵ 表示 $\hat{\boldsymbol{x}}$ 与 $\boldsymbol{x}^*$ 之间距离的上界：

$$\min\{\|\hat{\boldsymbol{x}}-\boldsymbol{x}^*\|_2,\|\hat{\boldsymbol{x}}+\boldsymbol{x}^*\|_2\}\leq\epsilon=\left(\frac{4}{v_0}\rho+\frac{4}{v_0}+1\right)\|\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-\boldsymbol{x}^*\|_2+\frac{4}{v_0}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}.$$

可以看出，测量值数量越多，速率 ρ 越低，进而失配误差也越小

$$\left(\frac{4}{v_0}\rho+\frac{4}{v_0}+1\right)\|\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-\boldsymbol{x}^*\|_2.$$

同时，由于对于给定的 ϵ，以下不等式成立：

$$\left(\frac{4}{v_0}\rho+\frac{4}{v_0}+1\right)\|\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-\boldsymbol{x}^*\|_2\leq\epsilon$$

经过一些简单计算可得

$$m\geq m_0\frac{16(\text{mismatch})^2}{(\epsilon v_0-(4+v_0)\,\text{mismatch})^2}\tag{52}$$

其中，我们用“失配”表示 $\|\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-\boldsymbol{x}^*\|_2$。因此，根据公式 (52)，如果失配项 $\|\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-\boldsymbol{x}^*\|_2$ 趋于零，则所需的 m 会更少，这与线性估计情形类似。

注释 IV.4. 与线性估计的情形类似，在条件$m\geq Cm_0$中指定 C 的具体取值后，我们可以推导出与定理 *IV.2* 中所述结果相当的定理，具体涉及定理 *IV.4*。与线性估计情形相似，对于具有重尾噪声 *[22]*,的对抗性相位恢复，采用非线性最小绝对偏差方法也存在类似的参数敏感性分析；这是一种自然推广，本文略去不述。

致谢

作者谨此感谢徐志强教授就相关理论进行的有益讨论。作者还要感谢审稿人提出的宝贵意见。

参考文献

[1] Y. Plan、R. Vershynin 和 E. Yudovina，“带几何约束的高维估计”，《信息与推断：IMA 杂志》，第 6 卷，第 1 期，第 1–40 页，2017 年。[2] S. Chatterjee、S. Chen 和 A. Banerjee，“广义 Dantzig 选择器：k-支撑范数的应用”，神经信息处理系统进展，2014 年。[3] Y. Plan 和 R. Vershynin，“非线性观测下的广义 Lasso”，IEEE 信息论汇刊，第 62 卷，第 3 期，第 1528–1537 页，2016 年。[4] A. Berk、Y. Plan 和 d. Yilmaz，“ l^1 最小化对参数选择的敏感性”，《信息与推断：IMA 杂志》，第 10 卷，第 2 期，第 397–453 页，2021 年。[5] A. Berk、Y. Plan 和 O. Yılmaz，“根据数据参数选择 Lasso 程序的最佳方案”，IEEE 信息论汇刊，第 68 卷，第 4 期，第 2573–2603 页，2022 年。[6] A. Berk、S. Brugiapaglia 和 T. Hoheisel，“重新审视 Lasso：变分分析视角及其在压缩感知中的应用”，SIAM 数据科学数学期刊，第 5 卷，第 4 期，第 1102–1129 页，2023 年。[7] S. Karmalkar 和 E. Price，“通过 l1 回归处理对抗性稀疏噪声的压缩感知”，载于第二届算法简洁性研讨会（SOSA 2019），第 69 卷。Schloss Dagstuhl – Leibniz 信息学中心，2019 年，第 19:1–19:19 页。[8] S. Foucart 和 H. Rauhut，“压缩感知入门”，载于《压缩感知数学导论》。施普林格，2013 年，第 1–39 页。[9] Y. Gordon，“关于 Milman 不等式及穿过 R^n 中网格逃逸的随机子空间”，载于《泛函分析的几何方面》。施普林格，1988 年，第 84–106 页。[10] R. Schneider 和 W. Weil，《随机几何与积分几何》。施普林格，2008 年。[11] S. S. Chen、D. L. Donoho 和 M. A. Saunders，“通过基追踪进行原子分解”，SIAM 科学计算期刊，第 20 卷，第 1 期，第 33–61 页，1998 年。[12] R. Tibshirani，“通过 Lasso 进行回归收缩与变量选择”，英国皇家统计学会会刊 B辑，第 58 卷，第 1 期，1996 年。[13] A. M. Bruckstein、D. L. Donoho 和 M. Elad，“从方程组的稀疏解到信号与图像的稀疏建模”，SIAM 评论，第 51 卷，第 1 期，第 34–81 页，2009 年。[14] S. Bahmani、B. R. Raj 和 P. T. Boufounos，“贪心稀疏约束优化”，机器学习研究期刊，第 14 卷，第 1 期，第 807–841 页，2013 年。[15] X.-T. Yuan、P. Li 和 T. Zhang，“用于稀疏约束优化的梯度硬阈值追踪”，载于国际机器学习会议（PMLR），2014 年，第 127–135 页。[16] S. Oymak、B. Recht 和 M. Soltanolkotabi，“线性逆问题中尖锐的时间-数据权衡”，IEEE 信息论汇刊，第 64 卷，第 6 期，第 4129–4158 页，2018 年。[17] S. Oymak 和 M. Soltanolkotabi，“从非线性观测中快速可靠地进行参数估计”，SIAM 优化期刊，第 27 卷，第 4 期，第 2276–2300 页，2017 年。[18] T. Blumensath 和 M. E. Davies，“压缩感知的迭代硬阈值”，《应用与计算调和分析》，第 27 卷，第 3 期，第 265–274 页，2009 年。[19] S. Bahmani 和 B. Raj，“用于 l_p约束最小二乘的投影梯度下降统一分析”，《应用与计算调和分析》，第 34 卷，第 3 期，第 366–378 页，2013 年。[20] Y. Zhong、Y. Shen 和 C. Li，“可压缩信号的 Frank-Wolfe 算法稳定性”，手稿，2024 年。[21] H. Xu、S. Li 和 J. Lin，“带对抗性稀疏噪声的低秩矩阵恢复”，《逆问题》，第 38 卷，第 3 期，文章编号 035001，2022 年。[22] G. Huang、S. Li 和 H. Xu，“通过非线性最小绝对偏差进行对抗性相位恢复”，arXiv，2023 年。 [在线]。链接：https://doi.org/10.48550/arXiv.2312.06190[23] S. Li、D. Liu 和 Y. Shen，“带稀疏约束的最小绝对偏差问题的自适应迭代硬阈值”，傅里叶分析与应用期刊，第 29 卷，第 5 期，2023 年。[24] D. Liu、S. Li 和 Y. Shen，“稀疏约束下采用投影次梯度法的单比特压缩感知”，IEEE 信息论汇刊，第 65 卷，第 10 期，第 6650–6663 页，2019 年。[25] X. Gao 和 J. Huang，“带 Lasso 的高维 LAD 回归渐近分析”，《思古统计学》，第 20 卷，第 1485–1506 页，2010 年。[26] L.-L. Sophie 和 Z. Laurent，“通过 Huber 准则和自适应 Lasso 惩罚进行稳健回归”，电子统计学期刊，第 5 卷，第 1015–1053 页，2011 年。

[27] H. Wang、G. Li 和 G. Jiang，《通过 lad-lasso 实现稳健回归收缩与一致变量选择》，《商业与经济统计杂志》，第 25 卷，第 347–355 页，2007 年。[28] L. Wang，《用于高维线性回归的 l_1 惩罚最小绝对偏差估计量》，《多元分析杂志》，第 120 卷，第 135–151 页，2013 年。[29] M. Soltanolkotabi，《从二次测量中恢复结构化信号：通过非凸优化突破样本复杂度障碍》，《IEEE 信息论汇刊》，第 65 卷，第 4 期，第 2374–2400 页，2019 年。[30] G. Wang、L. Zhang、G. Giannakis、M. Akcakaya 和 J. Chen，《通过截断幅度流进行稀疏相位恢复》，《IEEE 信号处理汇刊》，第 66 卷，第 2 期，第 479–491 页，2017 年。[31] H. Zhang、Y. Zhou、Y. Liang 和 Y. Chi，《相位恢复的非凸方法：重塑 Wirtinger 流与增量算法》，《机器学习研究杂志》，第 18 卷，2017 年。[32] M. Huang 和 Z. Xu，《非线性最小二乘法在相位恢复中的估计性能》，《IEEE 信息论汇刊》，第 66 卷，第 12 期，第 7967–7977 页，2020 年。[33] Y. Xia 和 Z. Xu，《基于振幅的模型在复数相位恢复中的性能》，《信息与推断：IMA 杂志》，第 13 卷，第 1 期，文章编号 iaad053，2024 年。[34] R. Hyder、V. Shah 和 C. e. a. Hegde，《利用生成式先验的交替相位投影梯度下降求解压缩相位恢复》，载于《ICASSP 2019-2019 IEEE 国际声学、语音与信号处理会议（ICASSP）》。IEEE，2019 年，第 7705–7709 页。[35] J. Cai、J. Li、X. Lu 和 J. You，《通过硬阈值追踪从无相位测量中恢复稀疏信号》，《应用与计算调和分析》，第 56 卷，第 367–390 页，2022 年。[36] G. Jagatap 和 C. Hegde，《从仅幅值测量中恢复结构化信号的样本高效算法》，《IEEE 信息论汇刊》，第 65 卷，第 7 期，第 4434–4456 页，2019 年。[37] J.A.Tropp，《从独立随机线性测量中凸恢复结构化信号》，载于《采样理论：复兴》。施普林格，2015 年，第 67–101 页。[38] Y. Liu、J. Wang、Y. Sun 和 Y. C. Eldar，《无稀疏性压缩感知及其在生成式先验中的应用》，《IEEE 信息论汇刊》，第 66 卷，第 12 期，第 7923–7943 页，2020 年。[39] R. Vershynin，《高维概率：数据科学中的导论与应用》。剑桥大学出版社，2020 年。[40] V. Roman，《高维估计：几何视角》，《采样理论：复兴：压缩感知及其他进展》，第 3–66 页，2015 年。[41] C. Liaw、A. Mehrabian、Y. Plan 和 R. Vershynin，《一种界定随机矩阵在几何集合上的偏差的简单工具》，《泛函分析的几何方面》，第 2169 卷，第 277–299 页，2017 年。[42] V. Voroninski 和 Z. Xu，《强限制等距性质及其在无相位压缩感知中的应用》，《应用与计算调和分析》，第 40 卷，第 2 期，第 386–395 页，2016 年。[43]M. Genzel 和 G. Kutyniok，《失配原理：大模型不确定性下的广义 Lasso》，arXiv 预印本 arXiv:1808.06329，2018 年。 [在线]。网址：https://arxiv.org/abs/1808.06329[44] S. Oymak、T. Christos 和 H. Babak，《广义 Lasso 的平方误差：精确分析》，载于 2013 年第 51 届 Allerton 通信、控制与计算会议。IEEE，2013 年，第 1002–1009 页。[45] Y. Plan 和 R. Vershynin，《通过随机超平面镶嵌进行降维》，《离散与计算几何》，第 51 卷，第 438–461 页，2014 年。

附录

引理 IV.1. *[39,* 高斯集中不等式*]* 设随机向量 $\boldsymbol{x} \sim \mathcal{N}(0, I_n)$ 和利普希茨函数 $g : \mathbb{R}^n \to \mathbb{R}$，其常数为 $\|g\|_L$：

$$|g(\boldsymbol{x}) - g(\boldsymbol{y})| \leq \|g\|_L \|\boldsymbol{x} - \boldsymbol{y}\|_2.$$

那么，对于每个 $t > 0$，事件 $|g(\boldsymbol{x}) - \mathbb{E}g(\boldsymbol{x})| \leq t$ 以至少 $1 - 2\exp\left(-\frac{Ct^2}{\|g\|_L^2}\right.$ 的概率成立 $\left.____\right)$。

与引理 III.1 类似，我们为 $\|A\boldsymbol{x}\|_1$ 给出以下集中不等式，它们在定理 IV.3 的估计中发挥着重要作用。

引理 IV.2. *[45, 引理 2.1]*设有一个有界子集 $\mathcal{T} \subset \mathbb{R}^n$ 以及独立随机向量 $\boldsymbol{a}_i \sim \mathcal{N}(0, I_n)$，$i = 1, 2, \ldots, m$。令

$$Z = \sup_{\boldsymbol{x}\in\mathcal{T}} \left| \frac{1}{m}\sum_{i=1}^{m} |\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle| - \sqrt{\frac{2}{\pi}}\|\boldsymbol{x}\|_2 \right|.$$

1) </style}有 $(Z) \leq {}^{4\omega}(\mathcal{T})\sqrt{m}$。

2) 对所有 $u > 0$，以下偏差不等式成立

$$\mathbb{P}\left[Z > \frac{4\omega(\mathcal{T})}{\sqrt{m}} + u \right] \leq 2\exp\left(-\frac{mu^2}{2\,\mathrm{rad}^2(\mathcal{T})} \right).$$

在定理 IV.3 的证明中，我们根据高斯随机矩阵（假设引理 III.1 成立）陈述引理 IV.2 的结果 [45, 注 2.2], 其中事件 Z 表示为

$$\sup_{\boldsymbol{x}\in\mathcal{T}} \left| \frac{1}{m}\|A\boldsymbol{x}\|_1 - \sqrt{\frac{2}{\pi}}\|\boldsymbol{x}\|_2 \right|.$$

引理 *III.5* 的证明：

证明。 对于任意固定的 $\boldsymbol{e} \in \mathbb{R}^m$，我们有

$$\mathbb{E}\sup_{\boldsymbol{x}\in\mathcal{T}} \langle \boldsymbol{x}, A^T\boldsymbol{e}\rangle = \|\boldsymbol{e}\|_2 \cdot \omega(\mathcal{T}),$$

这由高斯宽度的定义得出。接下来，我们设

$$g(A) := \sup_{\boldsymbol{x}\in\mathcal{T}} \langle \boldsymbol{x}, A^T\boldsymbol{e}\rangle.$$

对于任意矩阵 $A,\quad B \in \mathbb{R}^{m\times n}$，我们有

$$\begin{aligned} \left| \sup_{\boldsymbol{x}\in\mathcal{T}} \langle \boldsymbol{x}, A^T\boldsymbol{e}\rangle - \sup_{\boldsymbol{x}\in\mathcal{T}} \langle \boldsymbol{x}, B^T\boldsymbol{e}\rangle \right| &\leq \left| \sup_{\boldsymbol{x}\in\mathcal{T}} \langle (A-B)\boldsymbol{x}, \boldsymbol{e}\rangle \right| \\ &\leq \|\boldsymbol{e}\|_2 \|A - B\|_F, \end{aligned}$$

we 在最后一个不等式中使用 $\boldsymbol{x} \in \mathcal{T} \subset \mathcal{B}_2^n$ 这一事实。然后应用引理 IV.1，我们有 e

$$\sup_{\boldsymbol{x}\in\mathcal{T}} \langle \boldsymbol{x}, A^T\boldsymbol{e}\rangle \leq \mathbb{E}\sup_{\boldsymbol{x}\in\mathcal{T}} \langle \boldsymbol{x}, A^T\boldsymbol{e}\rangle + t$$

以至少为 $1 - 2\exp\left(-\frac{Ct^2}{\|g\|_L^2}\right)$ 的概率成立。选取 $t = \|\boldsymbol{e}\|_2 u$，其中 $u > 0$ 是任意的，我们得到

$$\sup_{\boldsymbol{x}\in\mathcal{T}} \langle \boldsymbol{x}, A^T\boldsymbol{e}\rangle \leq \|\boldsymbol{e}\|_2(\omega(\mathcal{T}) + u)$$

以至少 $1 - 2\exp(-Cu^2)$ 的概率。 □

定理 *IV.2* 的证明：

证明。 虽然该定理的证明与定理 IV.1 高度相似，但主要在界定 $\|A\boldsymbol{h}\|_2$ 的某些方面有所不同。

情形 1： $f(\boldsymbol{x}^*) \geq \eta$。由于 $\hat{\boldsymbol{x}}$ 是（1）的一个解，我们得到与（9）和（10）中相同的公式：

$$\|A\boldsymbol{h}\|_2^2 \leq \|A\boldsymbol{h}_2\|_2^2 + 2\langle A^T\boldsymbol{e}, \boldsymbol{h}_1\rangle,$$

and

$$\|A\boldsymbol{h}\|_2^2 \geq (\phi(m) - \phi(m_0))^2\|\boldsymbol{h}\|_2^2.$$

我们设 $\boldsymbol{h} = \hat{\boldsymbol{x}} - \boldsymbol{x}^* = \boldsymbol{h}_1 + \boldsymbol{h}_2$，其中 $\boldsymbol{h}_1 = \hat{\boldsymbol{x}} - \boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)$ 且 $\boldsymbol{h}_2 = \boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*) - \boldsymbol{x}^*$。与定理 IV.1 的证明不同，接下来我们使用引理 III.1 对 $\|A\boldsymbol{h}_2\|_2$ 进行界定。由引理 III.1 可知，以至少 $1 - 2\exp(-u^2)$ 的概率：

$$\left|\frac{\|A\boldsymbol{h}_2\|_2}{\phi(m)} - \|\boldsymbol{h}_2\|_2\right| \leq \delta\|\boldsymbol{h}_2\|_2$$

成立，由此可得

$$\|A\boldsymbol{h}_2\|_2^2 \leq (1+\delta)^2 m\|\boldsymbol{h}_2\|_2^2, \tag{53}$$

利用不等式 $\phi(m) \leq \sqrt{m}$。由于 $f(\boldsymbol{h}+\boldsymbol{x}^*) = f(\hat{\boldsymbol{x}}) \leq \eta \leq f(\boldsymbol{x}^*)$ 且 $f(\boldsymbol{h}_2+\boldsymbol{x}^*) = f(\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)) \leq \eta \leq f(\boldsymbol{x}^*)$，$\frac{\boldsymbol{h}}{\|\boldsymbol{h}\|_2}$ 和 $\frac{\boldsymbol{h}_2}{\|\boldsymbol{h}_2\|_2}$ 均属于集合 $\mathcal{C}_f(\boldsymbol{x}^*) \cap \mathcal{S}^{n-1}$。注意，随机过程 $\langle A^T\boldsymbol{e},\ \boldsymbol{h}_1\rangle = \langle A^T\boldsymbol{e},\ \boldsymbol{h}-\boldsymbol{h}_2\rangle$，且 $\boldsymbol{h}$ 和 $\boldsymbol{h}_2$ 均满足引理 III.5 中的条件，因此对于任意 $u > 0$,

$$\begin{aligned}
\langle A^T\boldsymbol{e}, \boldsymbol{h}_1\rangle &= \langle A^T\boldsymbol{e}, \boldsymbol{h} - \boldsymbol{h}_2\rangle \\
&\leq [\omega(\mathcal{C}_f(\boldsymbol{x}^*) \cap \mathcal{S}^{n-1}) + u]\|\boldsymbol{e}\|_2(\|\boldsymbol{h}\|_2 + \|\boldsymbol{h}_2\|_2) \\
&\leq \sqrt{m_0}\|\boldsymbol{e}\|_2(\|\boldsymbol{h}\|_2 + \|\boldsymbol{h}_2\|_2)
\end{aligned} \tag{54}$$

成立 以至少 $1 - 2\exp(-Cu^2)$ 的概率。结合式 (9)、(10)、(54) 和 (53)，可得

$$\begin{aligned}
(\phi(m) - \phi(m_0))^2\|\boldsymbol{h}\|_2^2 \leq \|A\boldsymbol{h}\|_2^2 &\leq \|A\boldsymbol{h}_2\|_2^2 + 2\langle A^T\boldsymbol{e}, \boldsymbol{h}_1\rangle \\
&\leq (1+\delta)^2 m\|\boldsymbol{h}_2\|_2^2 + 2\sqrt{m_0}\|\boldsymbol{e}\|_2(\|\boldsymbol{h}\|_2 + \|\boldsymbol{h}_2\|_2).
\end{aligned}$$

经计算可得

$$\begin{aligned}
(1-\rho)^2\|\boldsymbol{h}\|_2^2 &\leq \frac{(1+\delta)^2 m\|\boldsymbol{h}_2\|_2^2}{\phi^2(m)} + \frac{2\sqrt{m_0}\|\boldsymbol{e}\|_2(\|\boldsymbol{h}\|_2 + \|\boldsymbol{h}_2\|_2)}{\phi^2(m)} \\
&\leq 2(1+\delta)^2\|\boldsymbol{h}_2\|_2^2 + 4\rho(\|\boldsymbol{h}\|_2 + \|\boldsymbol{h}_2\|_2)\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}
\end{aligned}$$

其中，最后一个不等式是利用 $\phi^2(m) \geq \frac{1.0049}{2} m$ 得到的。因此，

$$\begin{aligned}
&\left((1-\rho)\|\boldsymbol{h}\|_2 - \frac{2\rho}{1-\rho}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}\right)^2 \\
\leq &\left(\sqrt{2}(1+\delta)\|\boldsymbol{h}_2\|_2 + \frac{2\sqrt{2}\rho}{3}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}\right)^2 + \left(\frac{2\rho}{1-\rho}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}\right)^2 - \left(\frac{2\sqrt{2}\rho}{3}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}\right)^2 \\
\leq &\left(\sqrt{2}(1+\delta)\|\boldsymbol{h}_2\|_2 + \left(\frac{4\sqrt{2}\rho}{3} + \frac{2\rho}{1-\rho}\right)\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}\right)^2.
\end{aligned}$$

对上述不等式两边开平方并重新排列各项，可得

$$\|\boldsymbol{h}\|_2 \leq \frac{\sqrt{2}(1+\delta)}{(1-\rho)}\|\boldsymbol{h}_2\|_2 + \left(\frac{4\sqrt{2}\rho}{3(1-\rho)} + \frac{2\rho}{(1-\rho)^2}\right)\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}.$$

情况 2： $0 < f(\boldsymbol{x}^*) < \eta$。由于 f 是绝对齐次的，我们有

$$f(\boldsymbol{x}^\eta) = \eta\frac{f(\boldsymbol{x}^*)}{f(\boldsymbol{x}^*)} = \eta.$$

其中 $\boldsymbol{x}^\eta = \frac{\eta}{f(\boldsymbol{x}^*)}\boldsymbol{x}^*$ 记 $\tilde{\boldsymbol{e}} = A\boldsymbol{x}^* - A\boldsymbol{x}^\eta + \boldsymbol{e}$，并且

$$\boldsymbol{h} = \hat{x} - \boldsymbol{x}^* = \boldsymbol{h}_1 - \boldsymbol{h}_2$$

其中 $\boldsymbol{h}_1 = \hat{\boldsymbol{x}} - \boldsymbol{x}^\eta$ 和 $\boldsymbol{h}_2 = \boldsymbol{x}^* - \boldsymbol{x}^\eta$，我们可以类似地得到与式 (14) 相同的关系：

$$\|A\boldsymbol{h}_1\|_2^2 \leq 2\langle A^T\tilde{\boldsymbol{e}}, \boldsymbol{h}_1\rangle. \tag{55}$$

与证明相似 定理 IV.1（情况 2）中，内积 $\langle A^T\tilde{\boldsymbol{e}},\ \boldsymbol{h}_1\rangle$ 和 $\|\tilde{\boldsymbol{e}}\|_2$ 也如（16）中所示有界 ：

$$\langle A^T\tilde{\boldsymbol{e}}, \boldsymbol{h}_1\rangle \leq \sqrt{m_0'}\|\tilde{\boldsymbol{e}}\|_2\|\boldsymbol{h}_1\|_2.$$

此外，注意到 $\frac{\boldsymbol{h}_2}{\|\boldsymbol{h}_2\|_2}$属于集合 $\mathcal{C}_f(\boldsymbol{x}^\eta) \cap \mathcal{S}^{n-1}$，因此在引理 III.1 中取 $\boldsymbol{x} = \frac{\boldsymbol{h}_2}{\|\boldsymbol{h}_2\|_2}$，可得

$$\begin{aligned}\|\tilde{\boldsymbol{e}}\|_2 &\leq \|A\boldsymbol{h}_2\|_2 + \|\boldsymbol{e}\|_2 \leq (\phi(m)(1+\delta))\,\|\boldsymbol{h}_2\|_2 + \|\boldsymbol{e}\|_2 \\ &\leq (1+\delta)\sqrt{m}\|\boldsymbol{h}_2\|_2 + \|\boldsymbol{e}\|_2 = (1+\delta)\sqrt{m}\|\boldsymbol{x}^\eta - \boldsymbol{x}^*\|_2 + \|\boldsymbol{e}\|_2,\end{aligned} \tag{56}$$

结合（55）、（15）、（16）和（56），我们得到

$$\begin{aligned}(\phi(m) - \phi(m_0'))^2\|\boldsymbol{h}_1\|_2^2 &\leq \|A\boldsymbol{h}_1\|_2^2 \leq 2\langle A^T\tilde{\boldsymbol{e}}, \boldsymbol{h}_1\rangle \leq 2\sqrt{m_0'}\|\tilde{\boldsymbol{e}}\|_2\|\boldsymbol{h}_1\|_2 \\ &\leq 2\sqrt{m_0'}\|\boldsymbol{h}_1\|_2\left((1+\delta)\sqrt{m}\|\boldsymbol{x}^\eta - \boldsymbol{x}^*\|_2 + \|\boldsymbol{e}\|_2\right),\end{aligned}$$

再结合引理 III.4 和 $\phi^2(m) \geq \frac{1.0049}{2}m$，可得——

$$\|\boldsymbol{h}_1\|_2 \leq \frac{4\rho'(1+\delta)}{(1-\rho)^2}\|\boldsymbol{x}^\eta - \boldsymbol{x}^*\|_2 + \frac{4\rho'}{(1-\rho)^2}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}. \tag{57}$$

最后，根据（57）和三角不等式，我们有

$$\begin{aligned}\|\boldsymbol{h}\|_2 &= \|\boldsymbol{h}_1 - \boldsymbol{h}_2\|_2 \\ &\leq \|\boldsymbol{h}_1\|_2 + \|\boldsymbol{x}^\eta - \boldsymbol{x}^*\|_2 \leq \left(\frac{4\rho'(1+\delta)}{(1-\rho')^2} + 1\right)\|\boldsymbol{x}^\eta - \boldsymbol{x}^*\|_2 + \frac{4\rho'}{(1-\rho')^2}\frac{\|\boldsymbol{e}\|_2}{\sqrt{m}}.\end{aligned}$$

□

定理 *IV.3* 的证明：**情形 1：** $f(\boldsymbol{x}^*) \geq \eta$。由于 $\hat{\boldsymbol{x}}$ 是 (2) 的一个解，我们有

$$\|A\hat{\boldsymbol{x}} - \boldsymbol{y}\|_1 \leq \|A\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*) - \boldsymbol{y}\|_1. \tag{58}$$

我们令

$$\boldsymbol{h} = \hat{\boldsymbol{x}} - \boldsymbol{x}^* = \boldsymbol{h}_1 + \boldsymbol{h}_2$$

其中

$$\boldsymbol{h}_1 = \hat{\boldsymbol{x}} - \boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*) \quad \text{and} \quad \boldsymbol{h}_2 = \boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*) - \boldsymbol{x}^*. \tag{59}$$

由（58）和（59）可得

$$\begin{aligned}\|A\boldsymbol{h}-\boldsymbol{e}\|_1 &= \|A\hat{\boldsymbol{x}}-A\boldsymbol{x}^*-\boldsymbol{e}\|_1\\ &= \|A\hat{\boldsymbol{x}}-\boldsymbol{y}\|_1\\ &\le \|A\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-\boldsymbol{y}\|_1\\ &= \|A\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-A\boldsymbol{x}^*-\boldsymbol{e}\|_1\\ &\le \|A\boldsymbol{h}_2\|_1+\|\boldsymbol{e}\|_1.\end{aligned}$$

因此,

$$\|A\boldsymbol{h}\|_1 \le \|A\boldsymbol{h}-\boldsymbol{e}\|_1+\|\boldsymbol{e}\|_1 \le \|A\boldsymbol{h}_2\|_1+2\|\boldsymbol{e}\|_1. \tag{60}$$

由于归一化误差 $\boldsymbol{h}/\|\boldsymbol{h}\|_2$ 属于集合 $\mathcal{T}=\mathcal{C}_f(\boldsymbol{x}^*)\cap\mathcal{S}^{n-1}$，通过在引理 IV.2 中取 $\boldsymbol{x}=\boldsymbol{h}/\|\boldsymbol{h}\|_2$，对于任意 $u>0$，我们有

$$\sup_{\boldsymbol{x}=\boldsymbol{h}/\|\boldsymbol{h}\|_2}\left|\frac{1}{m}\|A\boldsymbol{x}\|_1-\sqrt{\frac{2}{\pi}}\right| \le \frac{4\omega(\mathcal{T})}{\sqrt{m}}+u \le \gamma\sqrt{\frac{m_1}{m}}+u=\rho+u \tag{61}$$

以至少 $1-2\exp(-mu^2/2)$ 的概率成立。由 (61) 可得

$$\left(\sqrt{\frac{2}{\pi}}-\rho-u\right)\|\boldsymbol{h}\|_2 \le \frac{1}{m}\|A\boldsymbol{h}\|_1. \tag{62}$$

接下来，我们以类似的方式给出 $\|A\boldsymbol{h}_2\|_1$ 的一个界。由于归一化向量 $\boldsymbol{h}_2/\|\boldsymbol{h}_2\|_2$ 也属于集合 $\mathcal{T}$，因此在引理 IV.2 中取 $\boldsymbol{x}=\boldsymbol{h}_2/\|\boldsymbol{h}_2\|_2$，可得以下不等式以至少 $1-2\exp(-mu^2/2)$ 的概率成立:

$$\sup_{\boldsymbol{x}=\boldsymbol{h}_2/\|\boldsymbol{h}_2\|_2}\left|\frac{1}{m}\|A\boldsymbol{x}\|_1-\sqrt{\frac{2}{\pi}}\right| \le \frac{4\omega(\mathcal{T})}{\sqrt{m}}+u \le \rho+u,$$

由此可得

$$\frac{1}{m}\|A\boldsymbol{h}_2\|_1 \le \left(\sqrt{\frac{2}{\pi}}+\rho+u\right)\|\boldsymbol{h}_2\|_2 \tag{63}$$

结合（60）、（62）和（63），可得

$$\begin{aligned}\left(\sqrt{\frac{2}{\pi}}-\rho-u\right)\|\boldsymbol{h}\|_2 &\le \frac{1}{m}\|A\boldsymbol{h}\|_1\\ &\le \frac{1}{m}\|A\boldsymbol{h}_2\|_1+\frac{2}{m}\|\boldsymbol{e}\|_1\\ &\le \left(\sqrt{\frac{2}{\pi}}+\rho+u\right)\|\boldsymbol{h}_2\|_2+\frac{2}{m}\|\boldsymbol{e}\|_1.\end{aligned}$$

因此，通过计算可得

$$\|\hat{\boldsymbol{x}}-\boldsymbol{x}^*\|_2 \le \frac{\sqrt{\frac{2}{\pi}}+\rho+u}{\sqrt{\frac{2}{\pi}}-\rho-u}\|\boldsymbol{P}_{\mathcal{K}}(\boldsymbol{x}^*)-\boldsymbol{x}^*\|_2+\frac{2}{\sqrt{\frac{2}{\pi}}-\rho-u}\frac{\|\boldsymbol{e}\|_1}{m}.$$

情况 2： $0<f(\boldsymbol{x}^*)<\eta$。由于 f 是绝对齐次的，我们有

$$f(\boldsymbol{x}^\eta)=\eta\frac{f(\boldsymbol{x}^*)}{f(\boldsymbol{x}^*)}=\eta,$$

其中 $\boldsymbol{x}^{\eta} = {}_{f}{}^{\eta}{}_{(\boldsymbol{x}^*)}\boldsymbol{x}^*$。记 $\tilde{\boldsymbol{e}} = A\boldsymbol{x}^* - A\boldsymbol{x}^{\eta} + \boldsymbol{e}$，我们有

$$\boldsymbol{y} = A\boldsymbol{x}^* + \boldsymbol{e} = A\boldsymbol{x}^{\eta} + (A\boldsymbol{x}^* - A\boldsymbol{x}^{\eta} + \boldsymbol{e}) = A\boldsymbol{x}^{\eta} + \tilde{\boldsymbol{e}}.$$

设 $\boldsymbol{h} = \hat{\boldsymbol{x}} - \boldsymbol{x}^* = \boldsymbol{h}_1 - \boldsymbol{h}_2$，其中 $\boldsymbol{h}_1 = \hat{\boldsymbol{x}} - \boldsymbol{x}^{\eta}$ 且 $\boldsymbol{h}_2 = \boldsymbol{x}^* - \boldsymbol{x}^{\eta}$。由于 $\hat{\boldsymbol{x}}$ 是式 (2) 的解，以下不等式成立

$$\|A\hat{\boldsymbol{x}} - \boldsymbol{y}\|_1 \leq \|A\boldsymbol{x}^{\eta} - \boldsymbol{y}\|_1,$$

由此可得

$$\begin{aligned}\|A\boldsymbol{h}_1 - \tilde{\boldsymbol{e}}\|_1 &= \|A\hat{\boldsymbol{x}} - A\boldsymbol{x}^{\eta} - \tilde{\boldsymbol{e}}\|_1 \\ &= \|A\hat{\boldsymbol{x}} - \boldsymbol{y}\|_1 \\ &\leq \|A\boldsymbol{x}^{\eta} - \boldsymbol{y}\|_1 \\ &= \|\tilde{\boldsymbol{e}}\|_1.\end{aligned}$$

因此,

$$\|A\boldsymbol{h}_1\|_1 \leq \|A\boldsymbol{h}_1 - \tilde{\boldsymbol{e}}\|_1 + \|\tilde{\boldsymbol{e}}\|_1 \leq 2\|\tilde{\boldsymbol{e}}\|_1. \tag{64}$$

由于归一化误差 $\boldsymbol{h}_1/\|\boldsymbol{h}_1\|_2$ 属于集合 $\mathcal{T}' = \mathcal{C}_f(\boldsymbol{x}^{\eta}) \cap \mathcal{S}^{n-1}$，通过在引理 IV.2 中取 $\boldsymbol{x} = \boldsymbol{h}_1/\|\boldsymbol{h}_1\|_2$，我们得出对任意 $u > 0$

$$\sup_{\boldsymbol{x}=\boldsymbol{h}_1/\|\boldsymbol{h}_1\|_2} \left| \frac{1}{m}\|A\boldsymbol{x}\|_1 - \sqrt{\frac{2}{\pi}} \right| \leq \frac{4\omega(\mathcal{T})}{\sqrt{m}} + u \leq \gamma\sqrt{\frac{m_1'}{m}} + u = \rho' + u \tag{65}$$

以至少 $1 - 2\exp(-mu^2/2)$ 的概率成立。由式 (65) 可得

$$\left(\sqrt{\frac{2}{\pi}} - \rho' - u\right)\|\boldsymbol{h}_1\|_2 \leq \frac{1}{m}\|A\boldsymbol{h}_1\|_1. \tag{66}$$

类似地，$\boldsymbol{h}_2/\|\boldsymbol{h}_2\|_2$ 也属于集合 $\mathcal{T}'$，因此根据引理 IV.2，我们有

$$\frac{1}{m}\|A\boldsymbol{h}_2\|_1 \leq \left(\sqrt{\frac{2}{\pi}} + \frac{4\omega(\mathcal{T})}{\sqrt{m}} + u\right)\|\boldsymbol{h}_2\|_2 \leq \left(\sqrt{\frac{2}{\pi}} + \rho' + u\right)\|\boldsymbol{h}_2\|_2. \tag{67}$$

接下来，我们利用 (67) 对项 $\|\tilde{\boldsymbol{e}}\|_1 = \|A\boldsymbol{x}^* - A\boldsymbol{x}^{\eta} + \boldsymbol{e}\|_1$ 进行估计

$$\frac{1}{m}\|\tilde{\boldsymbol{e}}\|_1 \leq \frac{1}{m}\|A\boldsymbol{h}_2\|_1 + \frac{1}{m}\|\boldsymbol{e}\|_1 \leq \left(\sqrt{\frac{2}{\pi}} + \rho' + u\right)\|\boldsymbol{h}_2\|_2 + \frac{1}{m}\|\boldsymbol{e}\|_1. \tag{68}$$

结合式 (64)、(66)、(67) 和 (68)，我们推得

$$\begin{aligned}\left(\sqrt{\frac{2}{\pi}} - \rho' - u\right)\|\boldsymbol{h}_1\|_2 \leq& \frac{1}{m}\|A\boldsymbol{h}_1\|_1 \leq \frac{2}{m}\|\tilde{\boldsymbol{e}}\|_1 \\ \leq& 2\left(\sqrt{\frac{2}{\pi}} + \rho' + u\right)\|\boldsymbol{x}^{\eta} - \boldsymbol{x}^*\|_2 + \frac{2}{m}\|\boldsymbol{e}\|_1.\end{aligned} \tag{69}$$

最后，由式 (69) 可得

$$
\begin{aligned}
\|\boldsymbol{h}\|_2 &\leq \|\boldsymbol{h}_1\|_2 + \|\boldsymbol{x}^{\eta} - \boldsymbol{x}^*\|_2 \\
&\leq \frac{2\left(\sqrt{\frac{2}{\pi}} + \rho' + u\right)}{\sqrt{\frac{2}{\pi}} - \rho' - u}\left(\frac{\eta}{f(\boldsymbol{x}^*)} - 1\right)\|\boldsymbol{x}^*\|_2 + \frac{2}{\sqrt{\frac{2}{\pi}} - \rho' - u}\frac{\|\boldsymbol{e}\|_1}{m} + \left(\frac{\eta}{f(\boldsymbol{x}^*)} - 1\right)\|\boldsymbol{x}^*\|_2 \\
&= \frac{3\sqrt{\frac{2}{\pi}} + \rho' + u}{\sqrt{\frac{2}{\pi}} - \rho' - u}\left(\frac{\eta}{f(\boldsymbol{x}^*)} - 1\right)\|\boldsymbol{x}^*\|_2 + \frac{2}{\sqrt{\frac{2}{\pi}} - \rho' - u}\frac{\|\boldsymbol{e}\|_1}{m}.
\end{aligned}
$$