

## 线性逆问题中惩罚优化方法信号重建误差界研究

张欢<sup>\*①②</sup> 雷宏<sup>①</sup>

<sup>①</sup>(中国科学院电子学研究所 北京 100190)

<sup>②</sup>(中国科学院大学 北京 100049)

**摘要:** 惩罚优化问题常常用于在有噪声的条件下用较少的观测个数来求解线性逆问题。目前,对惩罚优化问题恢复误差的研究主要存在以下两点不足:一是对权重参数往往有要求;二是噪声的方向对误差的影响未知。针对这两个问题,该文研究了当存在有界噪声时,惩罚优化问题恢复的误差界。首先,该文从问题的几何出发,给定了一个几何条件。当这一条件满足时,就能够推导出惩罚优化问题恢复的一个明确的误差界。这个误差界保证了恢复的解是稳定的,也就是说,恢复误差不会超过观测误差的常数倍。同时,这一误差界对于任意的正权重参数都成立,并且揭示了恢复误差以及最优的权重选择与观测噪声的方向之间的联系。进一步地,当观测矩阵是一个高斯矩阵时,依据这一几何条件可以得到高概率稳定恢复所需的观测次数。仿真实验证明了理论结果的正确性。

**关键词:** 线性逆问题; 压缩感知; 稳定恢复; 惩罚优化问题; 权重选择

中图分类号: TN911.72

文献标识码: A

文章编号: 1009-5896(2019)12-2939-06

DOI: 10.11999/JEIT181125

## An Error Bound of Signal Recovery for Penalized Programs in Linear Inverse Problems

ZHANG Huan<sup>①②</sup> LEI Hong<sup>①</sup>

<sup>①</sup>(*Institute of Electronics, Chinese Academy of Sciences, Beijing 100190, China*)

<sup>②</sup>(*University of Chinese Academy of Sciences, Beijing 100049, China*)

**Abstract:** Penalized programs are widely used to solve linear inverse problems in the presence of noise. For now, the study of the performance of penalized programs has two disadvantages. First, the results have some limitations on the tradeoff parameters. Second, the effect of the direction of the noise is not clear. This paper studies the performance of penalized programs when bounded noise is presented. A geometry condition which is used to study the noise-free problems and constrained problems is provided. Under this condition, an explicit error bound which guarantees stable recovery (i.e., the recovery error is bounded by the observation noise up to some constant factor) is proposed. The results are different from many previous studies in two folds. First, the results provide an explicit bound for all positive tradeoff parameters, while many previous studies require that the tradeoff parameter is sufficiently large. Second, the results clear the role of the direction of the observation noise playing in the recovery error, and reveal the relationship between the optimal tradeoff parameters and the noise direction. Furthermore, if the sensing matrix has independent standard normal entries, the above geometry condition can be studied using Gaussian process theory, and the measurement number needed to guarantee stable recovery with high probability is obtained. Simulations are provided to verify the theoretical results.

**Key words:** Linear inverse problem; Compressed sensing; Stable recovery; Penalized program; Tradeoff

### 1 引言

线性逆问题指的是从线性观测中恢复真实信号。在很多实际应用中,人们常常遇到压缩采样的情况,例如信号处理<sup>[1-3]</sup>,图像处理<sup>[4]</sup>,机器学习<sup>[5]</sup>

等。在这种情况下,由于观测个数少于信号维度,因此,这一问题病态的。为了解决这一病态性,常常假设信号具有某些低维结构,然后通过求解优化问题来恢复真实信号。当信号先验或噪声幅度先验已知时,可以采用约束优化问题来恢复信号;而当信号先验和噪声幅度先验均未知时,一般采用惩罚优化问题来恢复信号。

收稿日期: 2018-12-06; 改回日期: 2019-04-01; 网络出版: 2019-05-28

\*通信作者: 张欢 zhanghuan13@mails.ucas.ac.cn

由于优化方法在实际问题中应用广泛,因此研究其恢复性能十分重要。许多学者在这方面做了大量工作<sup>[6-18]</sup>。例如,文献<sup>[6-8]</sup>研究了压缩感知中 $L_1$ 最小化问题的性能; Bayati等人<sup>[9]</sup>研究了压缩感知问题的相变; Chandrasekaran等人<sup>[10]</sup>研究了线性逆问题中约束优化问题的恢复误差界; Amelunxen等人<sup>[11]</sup>研究了无噪声条件下线性逆问题的紧相变现象; Oymak等人<sup>[12]</sup>研究了紧相变现象的普适性。而对惩罚优化问题, Negahban等人<sup>[13]</sup>研究了当结构函数满足可分解性质时,惩罚优化问题的恢复性能,所得结果要求权重参数足够大。Thrapoulidis等人<sup>[14]</sup>对恢复误差做了渐进的分析。文献<sup>[15-17]</sup>研究了结构噪声下惩罚优化问题的恢复性能。然而,这些结果往往存在以下不足:一是对权重参数的选择有要求(往往要求大于某个正阈值);二是噪声方向对恢复误差的影响未知。

针对上述问题,本文研究了当存在有界噪声时惩罚优化问题的恢复误差界。首先,根据问题的几何,给出了一个几何条件,这一条件曾用于分析无噪声和约束优化问题的性能<sup>[10]</sup>。然后,利用这一条件推导并获得了明确的恢复误差界,这个误差界依赖于权重参数的选择和观测噪声的方向,并且保证了恢复的解是稳定的。这一结果与前人的结果有如下不同:该误差界仅要求权重参数是正的,且与噪声方向有关,这就揭示了噪声方向对于恢复效果的影响,以及最优的权重选择和噪声方向之间的关系。除此之外,如果观测矩阵是随机的高斯矩阵,这个几何条件可以利用高斯过程理论来研究,然后得到了保证高概率稳定恢复所需的观测次数。最后,仿真实验证明了所得理论结果的正确性。

## 2 线性逆问题及惩罚优化方法

本节介绍线性逆问题的观测模型和惩罚优化方法。

### 2.1 线性逆问题

线性逆问题的观测模型为

$$\mathbf{y} = \mathbf{A}\mathbf{x}^* + \mathbf{z} \quad (1)$$

其中,  $\mathbf{y} \in \mathbf{R}^m$  是观测信号,  $\mathbf{A} \in \mathbf{R}^{m \times n}$  是观测矩阵,  $\mathbf{x}^* \in \mathbf{R}^n$  是真实信号,  $\mathbf{z} \in \mathbf{R}^m$  是观测噪声。这一问题的目标是已知  $\mathbf{y}$  和  $\mathbf{A}$  来恢复  $\mathbf{x}^*$ 。

### 2.2 惩罚优化问题

假设  $f: \mathbf{R}^n \rightarrow \overline{\mathbf{R}}$  是一个正常凸函数并且代表了真实信号  $\mathbf{x}^*$  的先验结果。当噪声幅度先验和信号先验  $f(\mathbf{x}^*)$  都未知时,可以通过求解式(2)的惩罚优化问题来恢复  $\mathbf{x}^*$

$$\min \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \cdot f(\mathbf{x}) \quad (2)$$

其中,  $0 < \lambda < \infty$  是一个权重参数。

## 3 准备知识

本节介绍一些后续分析将会用到的基础知识。

### 3.1 次梯度

对于一个正常凸函数  $h: \mathbf{R}^n \rightarrow \overline{\mathbf{R}}$ , 在  $\omega \in \mathbf{R}^n$  点的次梯度是

$$\partial h(\omega) = \{\mathbf{u} \in \mathbf{R}^n : h(\omega + \mathbf{d}) \geq h(\omega) + \langle \mathbf{u}, \mathbf{d} \rangle, \forall \mathbf{d} \in \mathbf{R}^n\} \quad (3)$$

### 3.2 下降锥和法锥

对于一个正常凸函数  $h: \mathbf{R}^n \rightarrow \overline{\mathbf{R}}$ , 在  $\omega \in \mathbf{R}^n$  点的下降锥是

$$D(h, \omega) = \{t \in \mathbf{R}^n : \exists a > 0, h(\omega + a \cdot t) \leq h(\omega)\} \quad (4)$$

函数  $h$  在  $\omega \in \mathbf{R}^n$  点的法锥是

$$N(h, \omega) = D(h, \omega)^\circ = \{\mathbf{u} \in \mathbf{R}^n : \langle \mathbf{u}, \mathbf{d} \rangle \leq 0 \forall \mathbf{d} \in D(h, \omega)\} \quad (5)$$

此外,假设次梯度  $\partial h(\omega)$  是非空、紧的且不包括原点,那么法锥和次梯度具有式(6)的关系<sup>[18]</sup>

$$N(h, \omega) = \{\mathbf{u} \in \mathbf{R}^n : \exists t \geq 0, \mathbf{u} \in t \cdot \partial h(\omega)\} \quad (6)$$

### 3.3 高斯宽度

对于任意集合  $T \subseteq \mathbf{R}^n$ , 其高斯宽度定义为

$$w(T) = \mathbf{E} \sup_{t \in T} \langle \mathbf{g}, t \rangle, \mathbf{g} \sim N(\mathbf{0}, \mathbf{I}_n) \quad (7)$$

### 3.4 符号说明

本节将文中用到的符号列表说明,见表1。

表1 本文符号说明

符号	说明
$\mathbf{y}$	观测信号
$\mathbf{A}$	观测矩阵
$\mathbf{x}^*$	真实信号
$\mathbf{z}$	观测噪声
$\varepsilon$	观测噪声 $\mathbf{z}$ 的幅度
$m$	观测个数(观测信号 $\mathbf{y}$ 的维度)
$n$	真实信号 $\mathbf{x}^*$ 的维度
$f$	真实信号 $\mathbf{x}^*$ 的结构函数
$\lambda$	惩罚优化问题中的权重调节参数
$\partial f(\mathbf{x}^*)$	函数 $f$ 在点 $\mathbf{x}^*$ 的次梯度
$D(f, \mathbf{x}^*)$	函数 $f$ 在点 $\mathbf{x}^*$ 的下降锥
$N(f, \mathbf{x}^*)$	函数 $f$ 在点 $\mathbf{x}^*$ 的法锥
$w(D(f, \mathbf{x}^*) \cap S^{n-1})$	下降锥 $D(f, \mathbf{x}^*)$ 的球面高斯宽度

## 4 主要结论

这一节陈述本文的主要结论。首先,对于无噪声线性逆问题,文献<sup>[6,7]</sup>首先研究了该问题的几

何, 即保证无噪声线性逆问题成功或失败的几何条件; 然后利用高斯过程理论或积分集合理论通过研究这一几何条件来获得观测个数与恢复性能的关系。在本文中, 借鉴这一研究思路, 同样地, 本文首先研究了保证惩罚优化问题稳定恢复的一个几何条件, 然后利用高斯过程理论研究这一几何条件得到保证高概率稳定恢复所需的观测个数。

首先给出保证惩罚优化问题能够实现稳定恢复的几何条件, 并且给出了具体的误差界。

**定理** 考虑惩罚优化问题来恢复真实信号  $\mathbf{x}^*$ , 记它的解为  $\hat{\mathbf{x}}$ 。假设以下几何条件满足

$$\overline{D}(f, \mathbf{x}^*) \cap \text{null}(\mathbf{A}) = \{\mathbf{0}\} \quad (8)$$

其中,  $\overline{D}(f, \mathbf{x}^*)$  代表下降锥  $D(f, \mathbf{x}^*)$  的闭包, 并且假设次梯度集合  $\partial f(\mathbf{x}^*)$  是非空、紧的且不包括原点。记  $\mathbf{z} = \mathbf{s} \cdot \varepsilon$ , 其中  $\mathbf{s}$  代表噪声的方向, 而  $\varepsilon = \|\mathbf{z}\|_2$  代表噪声的幅度。此外, 记  $\alpha = \lambda/\varepsilon$ 。那么有如式(9)的恢复误差界

$$\|\hat{\mathbf{x}} - \mathbf{x}^*\|_2 \leq \frac{2\text{dist}(\mathbf{A}^T \mathbf{s}, \alpha \cdot \partial f(\mathbf{x}^*))}{\sigma_{\min}(\mathbf{A}, C(\mathbf{s}, \alpha))} \cdot \varepsilon \quad (9)$$

其中,  $C(\mathbf{s}, \alpha)$  是一个锥, 定义为

$$C(\mathbf{s}, \alpha) = \{\mathbf{t} \in \mathbf{R}^n : \forall \boldsymbol{\omega} \in \partial f(\mathbf{x}^*), \\ \langle \mathbf{t}, \alpha \cdot \boldsymbol{\omega} - \mathbf{A}^T \mathbf{x} \rangle \leq 0\} \quad (10)$$

$\sigma_{\min}(\mathbf{A}, C(\mathbf{s}, \alpha))$  的定义是

$$\sigma_{\min}(\mathbf{A}, C(\mathbf{s}, \alpha)) = \inf_{\mathbf{d} \in C(\mathbf{s}, \alpha), \mathbf{d} \neq \mathbf{0}} \|\mathbf{A} \mathbf{d}\|_2 / \|\mathbf{d}\|_2 \quad (11)$$

$\text{dist}(\mathbf{A}^T \mathbf{s}, \alpha \cdot \partial f(\mathbf{x}^*))$  代表点  $\mathbf{A}^T \mathbf{s}$  到集合  $\alpha \cdot \partial f(\mathbf{x}^*)$  的距离

$$\text{dist}(\mathbf{A}^T \mathbf{s}, \alpha \cdot \partial f(\mathbf{x}^*)) = \inf_{\mathbf{d} \in \alpha \cdot \partial f(\mathbf{x}^*)} \|\mathbf{A}^T \mathbf{s} - \mathbf{d}\|_2 \quad (12)$$

**证明** 对任意的  $\varepsilon > 0$ , 考虑噪声  $\mathbf{z} \in \mathbf{R}^m$  并且满足  $\|\mathbf{z}\|_2 = \varepsilon$ 。记  $\mathbf{d} = \hat{\mathbf{x}} - \mathbf{x}^*$ 。由于  $\hat{\mathbf{x}}$  是惩罚优化问题的解

$$\frac{1}{2} \|\mathbf{A} \mathbf{d} - \mathbf{z}\|_2^2 + \alpha \varepsilon \cdot f(\mathbf{x}^* + \mathbf{d}) \leq \\ \frac{1}{2} \|\mathbf{z}\|_2^2 + \alpha \varepsilon \cdot f(\mathbf{x}^*) \quad (13)$$

移项、化简得

$$\frac{1}{2} \|\mathbf{A} \mathbf{d}\|_2^2 - \langle \mathbf{d}, \mathbf{A}^T \mathbf{z} \rangle \leq \alpha \varepsilon \cdot [f(\mathbf{x}^*) - f(\mathbf{x}^* + \mathbf{d})] \quad (14)$$

任取向量  $\boldsymbol{\omega} \in \partial f(\mathbf{x}^*)$ , 根据次梯度的定义

$$f(\mathbf{x}^*) - f(\mathbf{x}^* + \mathbf{d}) \leq -\langle \boldsymbol{\omega}, \mathbf{d} \rangle \quad (15)$$

结合式(14)和式(15), 有

$$\frac{1}{2} \|\mathbf{A} \mathbf{d}\|_2^2 - \langle \mathbf{d}, \mathbf{A}^T \mathbf{z} \rangle \geq -\alpha \varepsilon \cdot \langle \boldsymbol{\omega}, \mathbf{d} \rangle \quad (16)$$

即

$$\|\mathbf{A} \mathbf{d}\|_2^2 \leq 2\varepsilon \cdot \langle \mathbf{d}, \mathbf{A}^T \mathbf{s} - \alpha \boldsymbol{\omega} \rangle \quad (17)$$

其中,  $\mathbf{s} = \mathbf{z}/\varepsilon$  代表噪声的方向。接下来定义锥

$$C_1(\mathbf{s}, \alpha) = \{\mathbf{a} \in \mathbf{R}^n : \exists t > 0, \forall \boldsymbol{\omega} \in \partial f(\mathbf{x}^*), \\ \|\mathbf{A} \cdot t\mathbf{a}\|_2^2 \leq 2\varepsilon \cdot \langle t\mathbf{a}, \mathbf{A}^T \mathbf{s} - \alpha \boldsymbol{\omega} \rangle\} \\ = \{\mathbf{a} \in \mathbf{R}^n : \exists t > 0, \forall \boldsymbol{\omega} \in \partial f(\mathbf{x}^*), t \\ \|\mathbf{A} \mathbf{a}\|_2^2 \leq 2\varepsilon \cdot \langle \mathbf{a}, \mathbf{A}^T \mathbf{s} - \alpha \boldsymbol{\omega} \rangle\} \quad (18)$$

根据以上分析, 可以得到  $\mathbf{d} \in C_1(\mathbf{s}, \alpha)$ 。

目前为止, 以上证明还未用到条件  $\overline{D}(f, \mathbf{x}^*) \cap \text{null}(\mathbf{A}) = \{\mathbf{0}\}$ 。考虑到这一条件, 可以得到

$$C_1(\mathbf{s}, \alpha) \cap \text{null}(\mathbf{A}) = \{\mathbf{0}\} \quad (19)$$

用反证法证明这一结论。假设式(19)不成立, 那么必定可以找到一个非零的向量  $\mathbf{a}_0 \in C_1(\mathbf{s}, \alpha) \cap \text{null}(\mathbf{A})$ 。根据  $C_1(\mathbf{s}, \alpha)$  的定义, 对于任意的  $\boldsymbol{\omega} \in \partial f(\mathbf{x}^*)$ ,

$$\langle \mathbf{a}_0, \mathbf{A}^T \mathbf{s} - \alpha \boldsymbol{\omega} \rangle = \langle \mathbf{a}_0, -\alpha \boldsymbol{\omega} \rangle \geq t \\ \|\mathbf{A} \mathbf{a}_0\|_2^2 = 0 \quad (20)$$

这就意味着  $\langle \mathbf{a}_0, \boldsymbol{\omega} \rangle \leq 0$ 。由于次梯度集合  $\partial f(\mathbf{x}^*)$  非空、紧的, 且不包含原点, 根据次梯度、下降锥和法锥之间的关系, 有  $\mathbf{a}_0 \in \overline{D}(f, \mathbf{x}^*)$ 。这与条件  $\overline{D}(f, \mathbf{x}^*) \cap \text{null}(\mathbf{A}) = \{\mathbf{0}\}$  矛盾。因此, 式(19)得证。

接下来, 一方面, 利用式(19), 可以得到, 对于任意的  $\mathbf{a} \in C_1(\mathbf{s}, \alpha)$  且  $\mathbf{a} \neq \mathbf{0}$ , 必有  $\mathbf{a} \notin \text{null}(\mathbf{A})$ 。即式(21)成立

$$\mathbf{a} \in C_1(\mathbf{s}, \alpha), \mathbf{a} \neq \mathbf{0} \Rightarrow \langle \mathbf{a}, \alpha \boldsymbol{\omega} - \mathbf{A}^T \mathbf{s} \rangle < 0, \\ \forall \boldsymbol{\omega} \in \partial f(\mathbf{x}^*) \quad (21)$$

另一方面, 对于任意的  $\mathbf{a} \in \mathbf{R}^n$  满足  $\langle \mathbf{a}, \alpha \boldsymbol{\omega} - \mathbf{A}^T \mathbf{s} \rangle < 0, \forall \boldsymbol{\omega} \in \partial f(\mathbf{x}^*)$ , 那么一定可以找到一个足够小的  $t > 0$  使得

$$t \cdot \|\mathbf{A} \mathbf{a}\|_2^2 \leq 2\varepsilon \cdot \langle \mathbf{a}, \mathbf{A}^T \mathbf{s} - \alpha \boldsymbol{\omega} \rangle \quad (22)$$

这就说明  $\mathbf{a} \in C_1(\mathbf{s}, \alpha)$ 。言外之意, 有

$$\langle \mathbf{a}, \alpha \boldsymbol{\omega} - \mathbf{A}^T \mathbf{s} \rangle < 0, \forall \boldsymbol{\omega} \in \partial f(\mathbf{x}^*) \Rightarrow \mathbf{a} \in C_1(\mathbf{s}, \alpha) \quad (23)$$

结合式(21)和式(23), 可以知道

$$C_1(\mathbf{s}, \alpha) = \{\mathbf{0}\} \cup \{\mathbf{a} \in \mathbf{R}^n : \forall \boldsymbol{\omega} \in \partial f(\mathbf{x}^*), \\ \langle \mathbf{a}, \alpha \boldsymbol{\omega} - \mathbf{A}^T \mathbf{s} \rangle < 0\} \subseteq \{\mathbf{a} \in \mathbf{R}^n : \\ \forall \boldsymbol{\omega} \in \partial f(\mathbf{x}^*), \langle \mathbf{a}, \alpha \boldsymbol{\omega} - \mathbf{A}^T \mathbf{s} \rangle \leq 0\} \\ = C(\mathbf{s}, \alpha) \quad (24)$$

于是, 有  $\mathbf{d} \in C(\mathbf{s}, \alpha)$ 。根据  $\sigma_{\min}(\mathbf{A}, C(\mathbf{s}, \alpha))$  的定义和柯西-施瓦兹不等式

$$\sigma_{\min}^2(\mathbf{A}, C(\mathbf{s}, \alpha)) \cdot \|\mathbf{d}\|_2^2 \leq 2\varepsilon \cdot \|\mathbf{d}\| \\ \|\mathbf{A}^T \mathbf{s} - \alpha \boldsymbol{\omega}\|_2 \quad (25)$$

如果  $\sigma_{\min}(\mathbf{A}, C(\mathbf{s}, \alpha)) > 0$

$$\|\mathbf{d}\|_2 \leq \frac{2\|\mathbf{A}^T \mathbf{s} - \alpha \boldsymbol{\omega}\|_2}{\sigma_{\min}^2(\mathbf{A}, C(\mathbf{s}, \alpha))} \cdot \varepsilon \quad (26)$$

选择  $\boldsymbol{\omega} \in \partial f(\mathbf{x}^*)$  使得  $\|\mathbf{A}^T \mathbf{s} - \alpha \boldsymbol{\omega}\|_2$  最小, 就得到了定理1中的结论

$$\|\hat{\mathbf{x}} - \mathbf{x}^*\|_2 \leq \frac{2\text{dist}(\mathbf{A}^T \mathbf{s}, \alpha \cdot \partial f(\mathbf{x}^*))}{\sigma_{\min}^2(\mathbf{A}, C(\mathbf{s}, \alpha))} \cdot \varepsilon \quad (27)$$

因此, 为证定理1, 只需要证明  $\sigma_{\min}(\mathbf{A}, C(\mathbf{s}, \alpha)) > 0$ 。注意到锥  $C(\mathbf{s}, \alpha)$  是一个闭集, 所以只需要证明

$$C(\mathbf{s}, \alpha) \cap \text{null}(\mathbf{A}) = \{\mathbf{0}\} \quad (28)$$

为这一目的, 考虑任意  $\mathbf{a}_0 \in \text{null}(\mathbf{A})$  且  $\mathbf{a}_0 \neq \mathbf{0}$ 。由  $\mathbf{A}^T \mathbf{a}_0 = \mathbf{0}$  可知

$$\langle \mathbf{a}_0, \alpha \boldsymbol{\omega} - \mathbf{A}^T \mathbf{s} \rangle = \alpha \langle \mathbf{a}_0, \boldsymbol{\omega} \rangle \quad (29)$$

并且, 由于  $\overline{D}(f, \mathbf{x}^*) \cap \text{null}(\mathbf{A}) = \{\mathbf{0}\}$ , 必定有  $\mathbf{a}_0 \notin \overline{D}(f, \mathbf{x}^*)$ 。由次梯度、下降锥和法锥之间的联系可知, 必然存在某个  $\boldsymbol{\omega}_0 \in \partial f(\mathbf{x}^*)$  使得  $\langle \mathbf{a}_0, \boldsymbol{\omega}_0 \rangle > 0$ 。因此,  $\mathbf{a}_0 \notin C(\mathbf{s}, \alpha)$ , 故  $C(\mathbf{s}, \alpha) \cap \text{null}(\mathbf{A}) = \{\mathbf{0}\}$ 。

证毕

**说明1** 根据本文中的定理, 如果固定  $\alpha > 0$ , 选择权重参数  $\lambda = \alpha \cdot \varepsilon$ 。那么定理中的系数

$$\beta = \frac{2\text{dist}(\mathbf{A}^T \mathbf{s}, \alpha \cdot \partial f(\mathbf{x}^*))}{\sigma_{\min}^2(\mathbf{A}, C(\mathbf{s}, \alpha))} \quad (30)$$

就不依赖于噪声幅度  $\varepsilon$ 。因此, 当几何条件式(8)成立时, 恢复误差

$$\|\hat{\mathbf{x}} - \mathbf{x}^*\|_2 \leq \beta \cdot \varepsilon \quad (31)$$

且  $\beta$  与  $\varepsilon$  无关。这就说明如果权重参数正比于噪声幅度, 那么惩罚优化问题的解就是稳定的。并且, 注意到, 如果噪声是标准的高斯白噪声, 那么许多前人的工作, 如文献[6,16], 均得到了权重参数应当正比于标准差的结论。由于标准差控制着噪声的幅度大小, 因此这两种权重参数选择方式是类似的。

**说明2** 文献[13]研究了当结构函数  $f$  是一个范数且满足一定可分解性质时惩罚优化问题的恢复性能, 其结果要求权重参数足够大; 而本文定理仅要求  $f$  是一个合适的凸函数, 以及次梯度  $\partial f(\mathbf{x}^*)$  非空、紧和不包含原点, 且定理为任意的正权重参数都提供了一个恢复误差界。

**说明3** 一般而言, 不仅是噪声的幅度, 噪声的方向也会对恢复误差有一定的影响。定理就揭示了恢复误差与噪声方向之间的联系。除此之外, 如果认为最优权重参数应当使得恢复误差最小, 那么定理同样揭示了最优权重参数与噪声方向之间的联系。事实上, 最优的权重参数为

$$\lambda_{\text{opt}} = \alpha_{\text{opt}} \cdot \varepsilon \quad (32)$$

其中,  $\alpha_{\text{opt}}$  最小化式(30)中的常数  $\beta$ , 而  $\beta$  与噪声方向密切相关。

定理1中的结论依赖于几何条件式(8)。当观测矩阵是一个随机的高斯矩阵时, 这一几何条件可以用高斯过程理论来研究, 然后就可以得到保证高概率稳定恢复所需的观测个数。

**推论1** 假设观测矩阵  $\mathbf{A}$  的元素相互独立且满足标准高斯分布, 并且选择权重参数  $\lambda = \alpha \cdot \varepsilon$ , 其中  $\alpha > 0$  是一个固定的常数。对于任意的  $t \geq 0$ , 如果观测个数  $m$  满足

$$\mu_m \geq w(D(f, \mathbf{x}^*) \cap S^{n-1}) + t \quad (33)$$

那么惩罚优化问题能够实现稳定恢复的概率至少为  $1 - \exp(-t^2/2)$ 。其中,  $\mu_m = E\|\mathbf{g}\|_2, \mathbf{g} \sim N(\mathbf{0}, \mathbf{I})$ 。

**证明** 推论1的证明是标准的, 比如文献[10]。

**说明4** 注意到  $\mu_m$  有如下范围  $\sqrt{m-1} \leq \mu_m \leq \sqrt{m}$ 。因此, 推论1意味着如果观测个数

$$m \geq w^2(D(f, \mathbf{x}^*) \cap S^{n-1}) \quad (34)$$

那么惩罚优化问题就会以高概率实现稳定恢复。

## 5 仿真实验

本节利用一个仿真实验来验证上一节的理论结果。在这个实验中, 假设信号是稀疏的, 并且惩罚优化问题式(2)通过CVX<sup>[19]</sup>来求解。具体实验设置如下: 信号维度  $n = 64$ , 观测个数  $m$  和信号稀疏度  $s$  从1~64改变(以步进1), 并且噪声幅度  $\varepsilon$  取值  $[0.001, 0.01, 0.05 : 0.05 : 1.00]$ <sup>1)</sup>。对于每种选择, 按如下步骤进行实验:

(1)生成真实信号  $\mathbf{x}^* \in \mathbf{R}^n$ : 其中  $s$  个坐标满足标准高斯分布且相互独立, 另外  $n - s$  个设为0;

(2)生成观测矩阵  $\mathbf{A} \in \mathbf{R}^{m \times n}$ : 每个元素均服从标准高斯分布, 且相互独立;

(3)生成噪声向量  $\mathbf{z} \in \mathbf{R}^m$ : 首先生成  $\mathbf{z}_1 \in \mathbf{R}^m$ , 其元素服从标准高斯分布且相互独立, 令  $\mathbf{z} = \varepsilon \cdot \mathbf{z}_1 / \|\mathbf{z}_1\|_2$ ;

(4)获得观测信号  $\mathbf{y} \in \mathbf{R}^m$ :  $\mathbf{y} = \mathbf{A}\mathbf{x}^* + \mathbf{z}$ ;

(5)设  $\alpha = 1$ , 求解惩罚优化问题式(2), 并记录恢复误差  $\|\hat{\mathbf{x}} - \mathbf{x}^*\|_2$ ;

(6)以上过程重复20次, 并计算平均误差。

实验结果如图1所示。图1中的曲面代表平均恢复误差, 而竖直的平面代表推论1得到的稳定恢复的阈值, 即  $w^2(D(f, \mathbf{x}^*) \cap S^{n-1})$ 。可以看到, 当观测个数大于  $w^2(D(f, \mathbf{x}^*) \cap S^{n-1})$  时, 恢复误差与观

<sup>1)</sup> 意思为  $\varepsilon$  依次取以下值进行仿真实验: 0.001, 0.01, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 1.00。

测噪声的幅度成正比, 因此惩罚优化问题就能够实现稳定恢复(实际上, 仿真得到的保证稳定恢复的阈值相较本文推论1给出的阈值要稍大一些。这是因为, 推论1得到的阈值是一个紧阈值, 它相当于

同时实现50%概率稳定恢复和50%概率不稳定恢复的分界点。因此, 若要求稳定恢复的概率接近于1, 那么观测个数比较比推论1中的阈值大一些<sup>[7]</sup>。这就证明了上一节的理论结果是正确的。

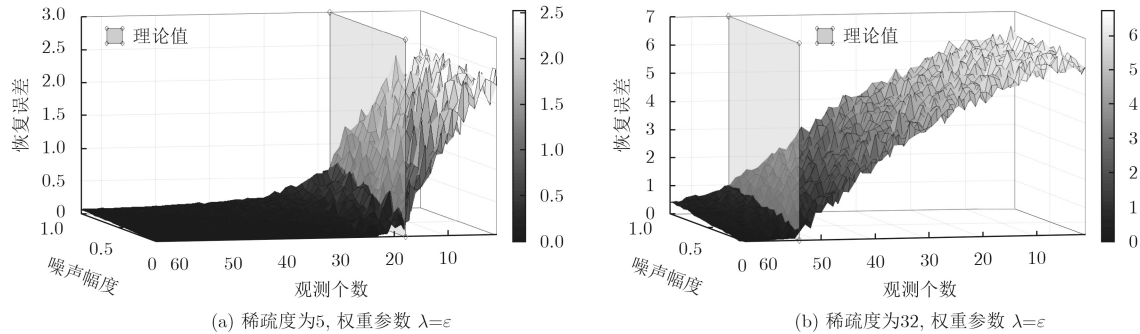


图1 数值仿真结果

## 6 结束语

本文研究了在线性逆问题中, 惩罚优化问题恢复真实信号时的恢复误差。本文分析依赖于一个几何条件, 它涉及信号结构函数的下降锥和观测矩阵的零空间。在这一条件下, 推导出了惩罚优化问题的一个明确的恢复误差界。这一误差界依赖于权重参数的选择和观测噪声的方向, 并且保证了解是稳定的。同时, 这一误差界适用于任意的正权重参数, 并且揭示了噪声方向对恢复误差的影响, 以及最优权重参数与噪声方向之间的联系。此外, 当观测矩阵是一个随机高斯矩阵时, 利用上述几何条件得到了保证高概率恢复所需的观测个数。仿真实验证明了理论结果的正确性。未来的研究方向包括不同观测矩阵、不同噪声模型下惩罚优化问题的恢复误差。

## 参考文献

- [1] ELDAR Y and KUTYNIOK G. Compressed Sensing: Theory and Applications [M]. Cambridge: Cambridge University Press, 2012: 210-268.
- [2] 游康勇, 杨立山, 刘玥良, 等. 基于稀疏贝叶斯学习的网格自适应多源定位[J]. 电子与信息学报, 2018, 40(9): 2150-2157. doi: 10.11999/JEIT171238.  
YOU Kangyong, YANG Lishan, LIU Yueliang, et al. Adaptive Grid Multiple Sources Localization Based on Sparse Bayesian Learning[J]. *Journal of Electronics & Information Technology*, 2018, 40(9): 2150-2157. doi: 10.11999/JEIT171238.
- [3] 王逸林, 马世龙, 王晋晋, 等. 基于稀疏重构的色噪声背景下未知线谱信号估计[J]. 电子与信息学报, 2018, 40(11): 2570-2577. doi: 10.11999/JEIT171040.  
WANG Yilin, MA Shilong, WANG Jinjin, et al. Estimation of unknown line spectrum under colored noise via sparse reconstruction[J]. *Journal of Electronics & Information Technology*, 2018, 40(11): 2570-2577. doi: 10.11999/JEIT171040.
- [4] STARCK J L, MURTAGH F, and FADILI J. Sparse Image and Signal Processing: Wavelets and Related Geometric Multiscale Analysis[M]. Cambridge: Cambridge University Press, 2016.
- [5] BISHOP C M. Pattern Recognition and Machine Learning[M]. New York: Springer, 2006.
- [6] CANDES E J, ROMBERG J K, and TAO T. Stable signal recovery from incomplete and inaccurate measurements[J]. *Communications on Pure and Applied Mathematics*, 2006, 59(8): 1207-1223. doi: 10.1002/cpa.2012.
- [7] CANDES E J and PLAN Y. Matrix completion with noise[J]. *Proceedings of the IEEE*, 2010, 98(6): 925-936. doi: 10.1109/JPROC.2009.2035722.
- [8] DONOHO D L, MALIKI A, and MONTANARI A. The noise-sensitivity phase transition in compressed sensing[J]. *IEEE Transactions on Information Theory*, 2011, 57(10): 6920-6941. doi: 10.1109/TIT.2011.2165823.
- [9] BAYATI M, LELARGE M, and MONTANARI A. Universality in polytope phase transitions and message passing algorithms[J]. *Annals of Applied Probability*, 2015, 25(2): 753-822. doi: 10.1214/14-AAP1010.
- [10] CHANDRASEKARAN V, RECHT B, PARRILO P A, et al. The convex geometry of linear inverse problems[J]. *Foundations of Computational Mathematics*, 2012, 12(6): 805-849. doi: 10.1007/s10208-012-9135-7.
- [11] AMELUNXEN D, LOTZ M, MCCOY M B, et al. Living on the edge: phase transitions in convex programs with random data[J]. *Information and Inference: A Journal of the IMA*, 2014, 3(3): 224-294. doi: 10.1093/imaiai/iau005.

- [12] OYMAK S and TROPP J A. Universality laws for randomized dimension reduction, with applications[J]. *Information and Interference*, 2017, 7(3): 337–446. doi: [10.1093/imaiai/iax011](https://doi.org/10.1093/imaiai/iax011).
- [13] NAGAHBAN S N, RAVIKUMAR P, WAINWRIGHT M J, et al. A unified framework for high-dimensional analysis of m-estimators with decomposable regularizers[J]. *Statistical Science*, 2012, 27(4): 538–557. doi: [10.1214/12-STS400](https://doi.org/10.1214/12-STS400).
- [14] THRAMPOULIDIS C, ABBASI E, and HASSIBI B. Precise high-dimensional error analysis of regularized m-estimators[C]. The 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton), Monticello, USA, 2015: 410–417. doi: [10.1109/MASS.1995.528223](https://doi.org/10.1109/MASS.1995.528223).
- [15] ZHANG Han, LIU Yulong, and LEI Hong. On the phase transition of corrupted sensing[C]. 2017 IEEE International Symposium on Information Theory, Aachen, Germany, 2017: 521–525. doi: [10.1109/ISIT.2017.8006582](https://doi.org/10.1109/ISIT.2017.8006582).
- [16] CHEN Jinchi and LIU Yulong. Corrupted sensing with sub-Gaussian measurements [C]. 2017 IEEE International Symposium on Information Theory, Aachen, Germany, 2017: 516–520. doi: [10.1109/ISIT.2017.8006581](https://doi.org/10.1109/ISIT.2017.8006581).
- [17] FOYGEL R and MACKEY L. Corrupted sensing: Novel guarantees for separating structured signals[J]. *IEEE on Transactions on Information Theory*, 2014, 60(2): 1223–1247. doi: [10.1109/TIT.2013.2293654](https://doi.org/10.1109/TIT.2013.2293654).
- [18] ROCKAFELLAR R T. Convex Analysis[M]. Princeton: Princeton University Press, 1970.
- [19] GRANT M and BOYD S. CVX: Matlab software for disciplined convex programming, version 2.1[OL]. <http://cvxr.com/cvx>, 2014.
- 张 欢: 男, 1991年生, 博士生, 研究方向为低维结构信号恢复理论及应用.
- 雷 宏: 男, 1963年生, 研究员, 博士生导师, 研究方向为电磁场与微波技术、信号处理理论与技术方面的研究.