

# 含噪稀疏信号重构的 $l_0$ 范数期望值最小化方法\*

王军华, 黄知涛, 周一宇

(国防科技大学 电子科学与工程学院, 湖南 长沙 410073)

**摘要:** 压缩感知理论是对信号压缩的同时进行感知的新理论, 而如何通过有限的测量值重构稀疏信号是压缩感知理论中的核心问题。针对测量值受噪声污染的含噪稀疏重构问题, 提出了近似  $l_0$  范数期望值最小化方法。该算法基本思想是将含噪稀疏重构问题转化为近似  $l_0$  范数期望值最小化问题, 并利用噪声的统计特征将随机最优化问题化简为常规的最优化问题, 然后采用最速下降法求解。数值仿真表明, 本文提出的方法具有更好的重构精度, 且计算量较小。

**关键词:** 压缩感知; 稀疏信号重构; 基追踪; 平滑  $l_0$  范数

**中图分类号:** TN911.72    **文献标志码:** A    **文章编号:** 1001-2486(2012)05-0045-04

## Sparse signal reconstruction with noise measurements based on expectation minimization of norm

WANG Junhua, HUANG Zhitao, ZHOU Yiyu

(College of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China)

**Abstract:** Compressed Sensing (CS) is a new framework for simultaneous sensing and compression, and how to recover sparse signal from limited measurements is the key problem in CS. A fast and stable method, called Expectation Minimization of approximate norm (abr. EML0), is proposed for sparse signal reconstruction with noisy measurements. The basic idea of the method is that sparse signal is recovered by minimizing the expectation of approximate norm, and then the expectation model by statistical character of noise is simplified so that the expectation model can be solved by the steepest descent method. Simulation results show that the proposed method provides better accuracy than existing methods at lower computational cost.

**Key words:** compressed sensing; sparse signal recovery; basis pursuit; smoothed norm

据奈奎斯特采样定理, 采样速率必须达到带宽的两倍以上, 才能完全重构信号。随着人们对信息需求量的增加, 携带信息的信号带宽越来越宽, 这无疑对采样速率和信号处理速度要求越来越高, 同时数据存储、传输以及处理都需要付出很大的代价。Donoho<sup>[1]</sup> 和 Candes<sup>[2]</sup> 提出的压缩感知 (Compressed Sensing or Compressive Sampling, CS) 为解决以上问题提供了思路, 而 CS 理论中的核心问题是如何用有限的测量值重构稀疏信号, 其本质是非欠定方程组求最稀疏解的问题, 该问题广泛应用于压缩感知的稀疏重构<sup>[1-2]</sup>、基于冗余字典的稀疏分解<sup>[3]</sup> 以及欠定盲源分离<sup>[4]</sup> 等领域。近年来, 稀疏重构问题受到越来越多学者的关注, 特别是针对含噪稀疏重构问题, 如何构造快速稳定的算法, 是稀疏信号重构问题的难点。

令信号  $z \in \mathbf{R}^n$  在某一组正交基或者紧框架  $\Psi$  的变换系数是稀疏的, 即  $z = \Psi x$ , 其中  $x \in \mathbf{R}^m$

是稀疏向量,  $\|x\|_0 = k \ll n$ ,  $k$  表示稀疏度。则稀疏信号的重构问题表示

$$y = \Phi \Psi x$$

其中  $\Phi$  是  $m \times n$  维观测矩阵,  $\Psi$  是  $n \times n$  维基矩阵, 向量  $y \in \mathbf{R}^m$  ( $m < n$ ) 为测量值。令  $A = \Phi \Psi$ , 则稀疏信号的重构问题等价于求解线性非欠定方程组  $y = Ax$  最稀疏解的问题, 也等价于  $l_0$  范数的最小化问题。但文献[5]指出,  $l_0$  范数最小化问题的求解需要组合搜索, 随着维数的增加难以实现。因此, 直接求解  $l_0$  范数最小化问题不可行, 需要寻求其他的方法。现有的方法主要有四类, 分别是 BP (Basis Pursuit) 算法<sup>[6-7]</sup>、贪婪算法<sup>[8-10]</sup>、IRLS 算法<sup>[11-13]</sup> 以及 SLO 算法<sup>[14-15]</sup>

其中最著名的是 BP 算法, 该算法将  $l_0$  范数最小化问题求解转化为  $l_1$  范数最小化问题求解。其基本思想是满足一定条件下  $l_1$  范数最小化与  $l_0$  范数最小化具有相同的稀疏解<sup>[21-22]</sup>。然而, BP

\* 收稿日期: 2012-01-12

基金项目: 国家自然科学基金资助项目 (61072120); 新世纪优秀人才支持计划资助项目 (NCET)

作者简介: 王军华 (1980—), 男, 江西景德镇人, 博士研究生, E-mail: skysword\_wjh@nudt.edu.cn;

周一宇 (通信作者), 男, 教授, 博士, 博士生导师, E-mail: zhouyiyu@shouhu.com

算法所需计算量较大,且重构稀疏向量时需要较多的测量值。贪婪类算法需要较少的计算量,但往往需要较多的观测值且估计精度不高。IRLS 算法将  $l_0$  范数最小化问题转化加权  $l_2$  范数的最小化问题,不同 IRLS 方法选择了不同的加权矩阵  $W$ 。IRLS 算法需要较少的测量,但收敛速度往往较慢,且容易收敛于局部最优解。SLO 算法则通过一类高斯函数近似  $l_0$  范数,从而把  $l_0$  范数最小化问题转化为近似  $l_0$  范数的最小化问题。SLO 算法计算速度快,且重构信号时仍然需要较少的测量值,但对噪声较为敏感。

在实际应用中,噪声往往是不可避免的,含噪稀疏重构问题表示为

$$y = Ax + \varepsilon \tag{1}$$

其中  $\varepsilon$  表示噪声。针对以上问题,学者们对 BP 算法加以改进,提出了相应的算法用于抑制噪声,最常见的方法有 LASSO 方法<sup>[16]</sup>和 DS<sup>[17]</sup>方法。LASSO 方法是  $l_1$  范数优化和  $l_2$  范数约束的最优化问题,而 DS 方法则是  $l_1$  范数优化和  $l_\infty$  范数约束的最优化问题。LASSO 算法和 DS 算法分别采用  $l_2$  和  $l_\infty$  范数作为约束条件,对噪声有一定的抑制作用,比 BP 算法更加稳健。但当信噪比较低时两种算法计算性能均欠佳,且和 BP 算法一样存在计算量大以及需要较多测量值的缺点。

针对含噪稀疏重构问题,本文提出了近似  $l_0$  范数期望值最小化方法 (Expectation Minimization of approximate  $l_0$  norm, EML0)。与已有的算法相比较,本文提出的方法具有更好的重构精度,且计算量较小。

### 1 EML0 算法基本思想

EML0 算法的基本思想是将含噪稀疏重构问题转化为近似  $l_0$  范数期望值最小化问题。首先将等式约束  $l_0$  范数最小化问题转化为含随机变量的非约束  $l_0$  范数最小化问题,然后利用高斯函数近似  $l_0$  范数并取期望值,建立近似  $l_0$  范数期望值最小化问题,最后利用噪声的统计特征将随机最优化问题转化为常规的最优化问题,然后用最速下降法求解。

含噪信号的稀疏重构问题(1),等价于等式约束的  $l_0$  范数最小化问题

$$\begin{cases} \min_x & \|x\|_0 \\ \text{s. t.} & Ax = y - \varepsilon \end{cases} \tag{2}$$

本文仅研究均值为 0,方差为  $\sigma^2 I_n$  的高斯白噪声。问题(2)中等式约束条件的  $x$  可以表示为

$$x = Fz + \tilde{x}$$

其中  $F = (F_1, F_2, \dots, F_m)^T$  是由  $A$  的零空间中任意一组基构成的  $n \times (n - m)$  维矩阵, $\tilde{x}$  是问题(2)中等式约束的特解。可以知道  $F$  是确定的,而特解  $\tilde{x}$  是与随机变量  $\varepsilon$  相关的,下面推导  $\tilde{x}$  关于  $\varepsilon$  的解析表达式。

问题(2)中的约束等式可以化为

$$(A_1 \ A_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = y - \varepsilon \Rightarrow A_1 x_1 + A_2 x_2 = y - \varepsilon$$

其中  $A = (A_1 \ A_2)$ ,  $x = (x_1 \ x_2)^T$ 。假设  $A_1$  可逆,令  $x_2 = 0$ ,则问题(2)中约束等式的一个特解可以表示为

$$\tilde{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} A^{-1}(y - \varepsilon) \\ 0 \end{pmatrix}$$

令  $A_1^{-1} = (B_1, B_2, \dots, B_m)^T$ ,则  $x$  可以表示为

$$x = \begin{cases} F_i z + B_i y - B_i \varepsilon & (i \leq m) \\ F_i z & (i > m) \end{cases} \tag{3}$$

由文献[13]给出的  $\|x\|_0$  的近似处理方法,令  $L$  取一个很小的值,可得

$$\|x\|_0 \approx n - \sum_{i=1}^n \exp(-x_i^2/2L^2) \tag{4}$$

可见,当  $L \rightarrow 0$  时,(4)式趋于相等。

将式(3)和式(4)代入问题(2)的目标函数中, $n$  是一个常数,对最优化结果没有影响,得到等价的无约束近似  $l_0$  最小化问题

$$\begin{aligned} \min_z f_L(z, \varepsilon) = & - \sum_{i=1}^m \exp[-(F_i z + B_i y - B_i \varepsilon)^2/2L^2] \\ & - \sum_{i=m+1}^n \exp[-(F_i z)^2/2L^2] \end{aligned} \tag{5}$$

式(5)是一个目标函数包含随机变量  $\varepsilon$  的最小化问题,因此我们取目标函数的期望值最小,得到近似  $l_0$  范数期望值问题

$$\begin{aligned} \min_z E[f_L(z, \varepsilon)] = & \min_z - \sum_{i=1}^m E\{\exp[-(F_i z + \\ & B_i y - B_i \varepsilon)^2/2L^2]\} - \sum_{i=m+1}^n \exp[-(F_i z)^2/2L^2] \end{aligned} \tag{6}$$

下面将利用噪声  $\varepsilon$  的统计特性简化近似  $l_0$  范数期望值最小化问题(6)。首先证明以下结论。

**结论 1** 若  $\beta$  是一满足高斯分布的随机变量,其均值为  $u$ ,方差为  $\sigma^2$ ,则

$$E[\exp(-\beta^2/2L^2)] = \frac{L}{\sqrt{\sigma^2 + L^2}} \exp[-u^2/2(\sigma^2 + L^2)]$$

**证明** 由  $\beta$  满足高斯分布可知

$$E[\exp(-\beta^2/2L^2)]$$

$$= \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} \exp(-\beta^2/2L^2) \exp(-(\beta-u)^2/2\sigma^2) d\beta$$

$$= \frac{L}{\sqrt{\sigma^2+L^2}} \exp[-u^2/2(\sigma^2+L^2)]$$

证明完毕。

易知当  $i \leq m$  时,  $x_i$  满足正态分布,其均值是  $F_i z + B_i y$ , 方差为  $\|B_i\|_2^2 \sigma^2$ , 利用结论 1 随机期望值优化模型(6)可以简化为

$$\min_z f_L(z) = - \sum_{i=1}^m \left[ \frac{L}{C_i} \exp[-(F_i z + B_i y)^2/2C_i^2] \right]$$

$$- \sum_{i=m+1}^n \exp(- (F_i z)^2/2L^2) \quad (7)$$

其中  $C_i = \sqrt{\|B_i\|_2^2 \sigma^2 + L^2}$  是不包括变量  $z$  的常数。信号的稀疏重构则通过求解最优化问题(7)得到。

随机期望值优化问题(7),是一个无约束优化问题,当参数  $L$  较小时,目标函数  $F_L(z)$  具有高度不平滑特征,使求解问题(7)时容易收敛于局部最优解。为了使最小化问题(7)能够尽量收敛于全局最优解,本文取参数  $L$  为一组下降序列

$$(L_1, L_2, \dots, L_J)$$

其中  $L_1$  为较大正常数,  $L_J$  为接近于零的正常数。然后,将参数  $L = L_{j-1}$  时得到(7)的最优解  $z_{L_{j-1}}$ , 作为参数为  $L = L_j$  时求解最小化问题(7)的初始值,从而使算法逐步靠近参数为  $L = L_J$  时的全局最优解。

当参数取  $L = L_j$  时最小化问题(7),是一个无约束最优化问题,用最速下降法求解得到  $\hat{z}$ 。式(3)的均值作为  $x$  的估计值  $\hat{x}$ , 即

$$\hat{x} = F \hat{z} + [A^{-1} y \quad 0]^T$$

## 2 EML0 算法

基于上节分析,本文提出了 EML0 算法,该算法的详细描述见算法 1。

算法注释如下:

一般取  $\delta_j = \gamma \delta_{j-1}, j=2, \dots, J, \gamma \in (0.5, 1)$ , 下面分别讨论参数  $\delta_1$  和  $\delta_j$  的选取。令  $\tilde{x} = \max_i |x_i^0|$ , 为了使算法能够较快地收敛,则令参数  $\delta_1$  满足

$$f_\delta(\tilde{x}) = \frac{2}{\pi} \arctan\left(\frac{\tilde{x}^2}{2\delta_1^2}\right) \leq \frac{1}{2} \Rightarrow \delta_1 \geq \frac{\tilde{x}}{\sqrt{2}}$$

为了节省计算量,本文取  $\delta_1 = \frac{\max_i |x_i^0|}{\sqrt{2}}$ 。当  $\delta_j$

$\rightarrow 0$  时,则  $F_{\delta_j}(x) \rightarrow \|x\|_0$ , 即  $\delta_j$  越小,则  $F_{\delta_j}(x)$  越能反映向量  $x$  的稀疏性,但同时噪声也越敏感。为了使对噪声不会太敏感,  $\delta_j$  的取值不宜过小。

## 算法 1: EML0 算法

算法输入: 矩阵  $A$ , 受噪声污染的测量值  $y$

初始化:

- (1) 选定初始值  $\hat{z}_0$ ;
- (2) 选择下降序列  $[L_1 L_2 \dots L_J]$ , 及终止准则  $\xi$ ;
- (3) 计算矩阵  $F$  和  $B$

算法迭代: for  $j=1, 2, \dots, J$

- (1) 令  $L = L_j, \xi = \xi_j$ ;
- (2) 用最速下降法求解最小化问题(7);
  - (a) 初始化  $\hat{z} = \hat{z}_{j-1}$ ;
  - (b) 计算  $f_L(z)$  在  $\hat{z}$  处梯度方向。  
 $d = \nabla f_L(\hat{z})$ , 若  $\|\nabla f_L(\hat{z})\|_2 \leq L\xi$ , 转到步骤(3);
  - (c) 搜索步长  $t$  使得  $F_{L_j}(\hat{z} - td) < F_{L_j}(\hat{z})$ ;
  - (d) 令  $\hat{z} = \hat{z} + td$ , 转步骤(b);
- (3) 令  $\hat{z}_j = \hat{z}$

输出结果:  $\hat{z} = \hat{z}_J, \hat{x} = F \hat{z} + [A^{-1} y \quad 0]^T$ ;

## 3 数值仿真

本节将通过仿真验证 EML0 算法的性能,并与 SLO<sup>[14]</sup>、LASSO<sup>[16]</sup>以及 DS<sup>[17]</sup>三种方法进行比较。仿真 1 比较了四种算法的计算时间;仿真 2 比较了噪声对四种算法性能的影响;仿真 3 比较了稀疏度对四种算法性能的影响。

下面所有数值仿真中,矩阵  $A$  是高斯随机矩阵,由  $\text{randn}(m, n)$  命令生成;稀疏信号  $x \in \mathbf{R}^n$  仅包含  $k$  个非零元素,且非零元素由均值为 0, 方差为 1 的高斯随机数构成;测量值向量通过  $y = Ax + \varepsilon$  生成,其中  $\varepsilon$  表示均值为 0, 方差为  $\sigma^2 I_n$  的高斯白噪声。重构信噪比定义为

$$RSNR = 10 \lg(\|x\|_2 / \|x - \hat{x}\|_2)$$

### 仿真 1 计算量的比较

本次仿真将通过计算时间来衡量算法的计算量。仿真环境如下: 软件是 MATLAB7. 9 和 Windows XP 操作系统;计算机配置为奔腾 2. 8G 处理器, 1G 内存。

表 1 分别比较了四种方法 200 次独立仿真的平均计算时间。可以看出, EML0 的计算时间略大于 SLO 方法, 但远小于 LASSO 和 DS 方法。

### 仿真 2 噪声对重构信噪比的影响

本次仿真将研究噪声均方差  $\sigma$  对算法性能的影响, 并与其他方法相比较。仿真参数设置为  $n = 256, m = 128, k = 40$ , 噪声均方差  $\sigma$  从 0. 02 增加到 0. 5。重构信噪比取 200 次实验的平均值。

图 1 比较了四种方法不同均方差时重构信噪比。由图 1 可以看出, LASSO 算法重构信噪比最低, 其次是 DS 算法和 SLO 算法, EML0 在不同噪

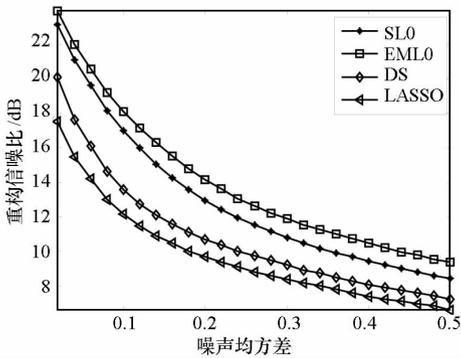


图 1 不同均方差时的重构信噪比

Fig. 1 RSNR comparison with various  $\sigma$

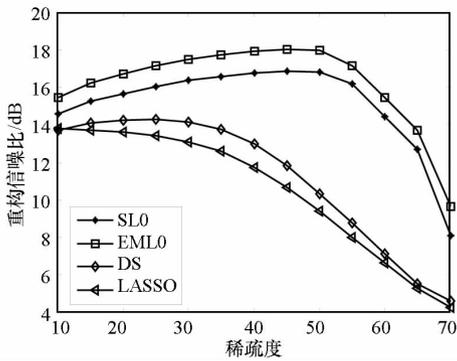


图 2 不同稀疏度时的重构信噪比

Fig. 2 RSNR comparison with various  $k$

声均方差条件下重构信噪比均高于 SL0 算法、DS 算法以及 LASSO 算法。

### 仿真 3 稀疏度对性能的影响

本次仿真将研究信号稀疏度  $k$  对算法性能的影响。取信号维数  $n = 256$ , 测量值数量  $m = 128$ , 噪声均方差  $\sigma = 0.1$ , 稀疏度  $k$  从 10 增加到 70。重构信噪比取 200 次实验的平均值。

图 2 比较了四种算法不同稀疏度  $k$  的重构信噪比。由图 2 可以看出, 四种算法中 EML0 算法的重构信噪比最高, LASSO 算法的重构信噪比最低, DS 算法的重构信噪比最接近 LASSO 算法。

表 1 EML0、SL0、DS 以及 LASSO 计算时间的比较

Tab.1 Compare computing time of EML0, SL0, DS and LASSO

$(n, m, k)$	EML0	SL0	DS	LASSO
(1000, 400, 100)	1.5150	0.9690	12.5922	25.8641
(1000, 200, 50)	0.6560	0.4850	8.4578	19.5531
(1000, 100, 30)	0.3440	0.2810	7.8469	18.5672

## 4 结论

针对含高斯白噪声的稀疏重构问题, 本文提出了 EML0 算法。该算法将含噪稀疏重构问题转化为近似  $l_0$  范数期望值最小化问题, 并利用噪声的统计特征将随机最优化问题化简为常规的最优化问题,

然后采用最速下降法求解。由于利用了噪声的统计特征, 提高了算法对噪声的适应能力。数值试验表明该算法能够提高估计精度, 且计算量较小。

## 参考文献 (References)

- [1] Donoho D L. Compressed sensing [J]. IEEE Transactions. on Information Theory, 2006, 52(4): 1289 - 1306.
- [2] Candes E. Compressive sampling [C]//Proceedings of the International Congress of Mathematics, Madrid, Spain. 2006, 3, 1433 - 1452.
- [3] Donoho D L, Elad M, Temlyakov M. Stable recovery of sparse over-complete representations in the presence of noise [J]. IEEE Transactions. On Information Theory, 2006, 52(1): 6 - 18.
- [4] Boll P, Zibulevsky M. Underdetermined blind source separation using sparse representations [J]. Signal Processing, 2001, 81: 2353 - 2362.
- [5] Donoho D L. For most large underdetermined systems of linear equations: the minimal  $L_1$  norm solution is also the sparsest solution [J]. Comm. Pure Appl. Math., 2006, 59(6): 797 - 829.
- [6] Baraniuk R. A lecture on compressed sensing [J]. IEEE Signal Processing Magazine, 2007, 24(4): 118 - 124.
- [7] Candès E, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information [J]. IEEE Transactions. Inf. Theory, 2006, 52: 489 - 509.
- [8] Tropp J, Gilbert A. Signal recovery from random measurements via orthogonal matching pursuit [J]. IEEE Transactions. Inf. Theory, 2007, 53(12): 4655 - 4666.
- [9] Needell D, Vershynin R. Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit [J]. IEEE J. Sel. Topics Signal Process, 2010, 4(2): 310 - 316.
- [10] Dai W, Milenkovic O. Subspace pursuit for compressive sensing signal reconstruction [J]. IEEE Transactions on Signal Processing, 2009, 55(5): 2230 - 2249.
- [11] Mourad N, Reilly J P. Minimizing non-convex functions for sparse vector reconstruction [J]. IEEE Transactions on Signal Processing, 2010, 58(7): 3485 - 3496.
- [12] Gorodnitsky I F, Rao B D. Sparse signal reconstructions from limited data using FOCUS: A re-weighted minimum norm algorithm [J]. IEEE Transactions on Signal Processing, 1997, 45(3): 600 - 616.
- [13] Chartrand R, Yin W. Iteratively re-weighted algorithms for compressed sensing [J]. Proc. Acoustics, Speech and Signal Processing, 2008, 3869 - 3872.
- [14] Mohimani H, Zadeh M, Jutten C. A fast approach for over-complete sparse decomposition based on smoothed  $L_0$  norm [J]. IEEE Transactions on Signal Processing, 2009, 57(1): 289 - 301.
- [15] Hyder M, Mahata K. An improved smoothed  $L_0$  approximation algorithm for sparse representation [J]. IEEE Transactions on Signal Processing, 2010, 58(4): 2194 - 2205.
- [16] Tibshirani R. Regression shrinkage and selection via the lasso [J]. J. Roy. Statist. Soc. Ser, 1996, 58: 267 - 288.
- [17] Tropp J. Just relax: convex programming methods for identifying sparse signals in noise [J]. IEEE Transactions. Inf. Theory, 2006, 52(3): 1030 - 1051.