压缩感知问题的目标罚函数交替随机搜索方法

蒋 敏[1] 孟志青[1] 沈 瑞[2]

(浙江工业大学管理学院 杭州 310023)[1] (浙江工业大学经济学院 杭州 310023)[2]

摘 要 首先将压缩感知优化问题等价定义为双凸优化问题,证明了这个等价双凸优化问题的最优解也是压缩感知优化问题的最优解,然后定义了它的一个具有 2 阶以上的光滑性的目标罚函数及对应的交替子问题,给出了一个交替求解子问题迭代算法,理论上证明了所提出的交替算法的收敛性定理,导出了压缩感知的最优解显示表达式,设计了一种对一类特定的压缩感知问题有效的交替随机搜索算法。该方法为研究和解决实际的压缩感知问题提供了一种新的设计思路。

关键词 压缩感知,等价表示,稀疏优化,目标罚函数,交替随机搜索算法

中图法分类号 TP391.4 **文献标识码** A

Alternate Random Search Algorithm of Objective Penalty Function for Compressed Sensing Problem

JIANG Min[1] MENG Zhi-qing[1] SHEN Rui[2]

(School of Management,Zhejiang University of Technology,Hangzhou 310023,China)[1]

(School of Economics,Zhejiang University of Technology,Hangzhou 310023,China)[2]

Abstract The compressed sensing optimization problem was defined as a biconvex optimization problem. It is proved that the optimal solution of the equivalent biconvex optimization problem is also the optimal solution of the compressed sensing optimization problem. Then a smooth objective penalty function and its corresponding alternating sub-problem were defined. An iterative algorithm for solving the sub-problem was given. The convergence theorem of alternating algorithm was proved theoretically. The expression of the optimal solution for compression perception was derived. An alternating random search algorithm was designed,which is effective for a specific type of compressed sensing problem. This method provides a new design idea for studying and solving the actual compressed sensing problem.

Keywords Compressive sensing, Equivalent representation, Sparse optimization, Object penalty function, Alternating random search algorithm

1 引言

Candès 和 Tao[1] 于 2006 年证明了压缩感知优化问题的正则化表示 l_1 范数在一定条件下可以用 l_0 范数求解,即信号 x 恢复问题 P $\min\limits_x \|x\|_0$ s. t. $Ax=b$ 由求解(P_p) $\min\limits_x \|x\|_p^p$ s. t. $Ax=b$ 替代,其中 $p=1$ 方便转换成无约束问题 $\min\limits_x \lambda\|x\|_p^p+\|Ax-b\|_2^2$,甚至可用线性规划求解,但是存在无解的情形,例如恢复问题[2] $A=\begin{pmatrix}2&1&1\\1&1&2\end{pmatrix}$,$b=\begin{pmatrix}1\\1\end{pmatrix}$的最优信号为 $x=(0,1,0)^{\mathrm{T}}$,但是(P_1)的解为 $x=(1,0,1)^{\mathrm{T}}$。Candès 等[2] 于 2008 年提出了加权 l_1 范数恢复问题:(WP)$\min\limits_x\sum\limits_{i=1}^n w_i|x_i|$ s. t. $Ax=b$,以改善(P_1)存在的不可恢复情形,进一步,Chartrand 等[3] 同样针对(P_p)的类似问题考虑加权算法($0<p<1$)。Mohimani 等[4] 提出一种光滑松弛问题$\min\limits_x\sum\limits_{i=1}^n(1-\exp(\frac{-x_i^2}{2\sigma^2}))$ s. t. $Ax=b$,完美恢复了问题 P。Foucart 等[5] 给出了几种替代(P_1)的近似替代求解方法,Pant 等[6] 推广了文献[3]的方法,使用$\min\limits_x \lambda\sum\limits_{i=1}^n(x_i^2+\varepsilon^2)^{p/2}+\|Ax-b\|_2^2$ 替代求解(P_p),由于在 0 点的局部凸性,实验结果具有较好的恢复效果。Wang 等[7] 特别讨论了 l_p($0<p<1$)范数的恢复性质,针对文献[3]的加权情形进行了对比实验,表明了其比加权 l_1 范数恢复问题有更好的效果。Zhu 等[8] 构造了增广拉格朗日函数松弛问题逼近 l_1 范数恢复问题 P_1,其算法具有全局收敛性。文婷婷等[9] 针对高斯光滑松弛问题提出了基于拟牛顿法的压缩感知重构零范数平滑算法。杜卓明等[10] 提出一个正则逼近函数$\sum\limits_{i=1}^n[\frac{2}{\pi}\arctan(-\exp(\beta x_i^2+1)]^2$ 的光滑正则化压缩感知信号重构方法,实验表明其优于前面提出的几种方法。

罚函数是求解约束优化的一种重要方法,它是将约束优化问题转化成无约束优化问题的方法,对于简化约束优化问题起到关键作用。该方法由于使用简单、便于掌握,受到许多工程优化研究者和应用者的青睐。压缩感知的 p 模问题的正则化表示本质上可以看作一种精确罚函数的形式[1]。2013

本文受浙江省自然科学基金项目(LY18A010031),国家自然科学基金项目(11871434)资助。

蒋 敏(1976-),女,教授,博士生导师,主要研究方向为供应链管理、非线性优化理论、机器学习、数据挖掘;**孟志青**(1962-),男,教授,博士生导师,主要研究方向为供应链管理、非线性优化理论、机器学习、数据挖掘,E-mail:mengzhiqing@zjut.edu.cn(通信作者);**沈 瑞**(1969-),女,博士,讲师,主要研究方向为供应链管理、非线性优化理论、机器学习、数据挖掘。

年，Meng 等[11]发展了一种新的精确目标罚函数来求解约束优化问题。鉴于此思想，我们研究了一种压缩感知优化问题的等价表示及其目标罚函数方法[13]。本文提出了一个 0 模(P)等价的双凸优化问题，定义它的目标罚函数，目标罚函数也为双凸函数，基于此，提出了一个交替求解算法，并给出了收敛性证明，进一步得到了(P)的最优解解析表达式；在此基础上构造了一个求解压缩感知问题的随机求解算法。该方法简单、容易实现，避开了求解优化问题的复杂性，表明了通过目标罚函数构造压缩问题的算法有效性。

2 目标罚函数表示及其算法

设变量 $x\in R^n$，$m\times n$ 矩阵 A，向量 $b\in R^m$，压缩感知问题为：

$$(P_0)\ \min_x \|x\|_0\ \text{ s.t. } Ax=b$$

学术界几乎都采用的松弛问题表示(正则化问题)[2]为：

$$(P_1)\ \min_x \|Ax-b\|_2^2+\lambda\|x\|_p^p$$

其中，参数 $\lambda,p>0$。在一定条件下可以求解问题 P_1 得到问题 P_0 的解，由于问题 P_1 不是凸问题，很难求得原问题的全局最优解。

下面提出问题 P_0 的另一种双凸优化问题 D_0，其表示为：

$$\min_{x,y,z} f(x,y;z)=\sum_{i=1}^n y_i$$

$$\text{s.t. } Ax=b,x_iz_i=0,y_iz_i=0,y_i+z_i=1,i=1,2,\cdots,n$$

问题 D_0 是一个关于$(x,y;z)$的双线性优化问题(双凸优化)。设可行解集合为：

$$S=\{(x,y,z)\mid Ax=b,x_iz_i=0,y_iz_i=0,y_i+z_i=1,i=1,2,\cdots,n\} \tag{1}$$

根据文献[12]，我们有问题 D_0 的局部最优解(partial optimum)定义：对于给定的$(x^*,y^*,z^*)\in S$，有 $f(x^*,y^*;z^*)\leqslant f(x,y;z^*)$，$\forall(x,y;z^*)\in S$，$f(x^*,y^*;z^*)\leqslant f(x^*,y^*;z)$，$\forall(x^*,y^*;z)\in S$。

定理 1 (1)若问题 D_0 存在可行解，则问题 P_0 也存在可行解。

(2)若$(x^*,y^*;z^*)$是问题 D_0 的最优解，则 x^* 是问题 P_0 的最优解，且 $\|x^*\|_0=\sum_{i=1}^n y_i^*$。

定理 1 说明通过求解问题 D_0 的最优解就可以得到压缩感知问题 P_0 的最优解。反之，若 x^* 是问题 P_0 的最优解，若 $x_i^*\neq0$，则令 $y_i^*=1$，否则，令 $y_i^*=0$，得到(x^*,y^*)是问题 D_0 的最优解。因此，本质上，问题 P_0 与问题 D_0 是等价的。

为了简化约束个数，问题 D_0 的另一种优化问题 D_1 表示为：

$$\min_{x,y,z} f(x,y;z)=\sum_{i=1}^n y_i$$

$$\text{s.t. } Ax=b,\sum_{i=1}^n[(x_i^2+y_i^2)z_i^2+(y_i+z_i-1)^2]=0$$

显然，问题 D_0 等价于问题 D_1。设 $A_j=(a_{j1},a_{j2},\cdots,a_{jn})$，$j=1,2,\cdots,m$，定义一种目标罚函数：

$$F(x,y;z;M,\rho)=\max\{\sum_{i=1}^n y_i-M,0\}^2+\rho\sum_{i=1}^n[(x_i^2+y_i^2)z_i^2+(y_i+z_i-1)^2]+\rho\sum_{j=1}^m(\max\{A_jx-b_j,0\}^2\max\{-A_jx+b_j,0\}^2) \tag{2}$$

目标罚函数 $F(x,y;z;M,\rho)$分别关于变量(x,y)(固定z)和 z(固定(x,y))是凸函数，是关于$(x,y;z)$的一个双凸函数。式(2)是问题 D_0 或问题 D_1 的目标罚函数。设：

$$C(x,y,z)=\sum_{i=1}^n[(x_i^2+y_i^2)z_i^2+(y_i+z_i-1)^2]+\sum_{j=1}^m(\max\{A_jx-b_j,0\}^2+\max\{-A_jx+b_j,0\}^2) \tag{3}$$

设函数：$g_j(x)=A_jx-b_j(j=1,2,\cdots,m)$，$h(x,y,z)=\sum_{i=1}^n[(x_i^2+y_i^2)z_i^2+(y_i+z_i-1)^2]$，那么问题 D_1 的 KKT 条件可以写为：存在不全为 0 的常数 $\alpha_1,\alpha_2,\cdots,\alpha_m,\alpha_{m+1}$，使得

$$\nabla f(x,y;z)+\sum_{j=1}^m\alpha_j\nabla g_j(x)+\alpha_{m+1}\nabla h(x,y;z)=0$$

即等价于下面的方程组：

$$\sum_{j=1}^m\alpha_jA_j^{\mathrm{T}}+\alpha_{m+1}\nabla_xh(x,y;z)=0$$

$$\nabla_yf(x,y;z)+\alpha_{m+1}\nabla_yh(x,y;z)=0$$

$$\alpha_{m+1}\nabla_zh(x,y;z)=0$$

定义 2 个无约束凸优化子问题：

$$D_0(M,\rho)_z:\ \min_{(x,y)\in R^n\times R^n} F(x,y;z;M,\rho)$$

$$D_0(M,\rho)_{(x,y)}:\min_{z\in R^n} F(x,y;z;M,\rho)$$

其中问题 $D_0(M,\rho)_z$ 固定变量 z 和参数(M,ρ)，求最优的(x,y)；问题 $D_0(M,\rho)_{(x,y)}$ 固定变量(x,y)和参数(M,ρ)，求最优的 z。对 $F(x,y;z;M,\rho)$求导$\frac{\partial F}{\partial z_i}=0$，得到：

$$z_i^{\#}=\frac{1-y_i}{1+x_i^2+y_i^2},i=1,2,\cdots,n \tag{4}$$

是问题 $D_0(M,\rho)_z$ 的最优解。因此，若求问题 D_0 的最优解，只要求出 $D_0(M,\rho)_z$ 的最优解即可。我们有下面的定理。

定理 2 设(x^*,y^*,z^*)是问题 D_0 的最优解，给定 $M<0$，设(x_M^*,y_M^*)是问题 $D_0(M,\rho)_{z_M^*}$ 的最优解，z_M^* 是问题 $D_0(M,\rho)_{(x_M^*,y_M^*)}$ 的最优解，如果(x_M^*,y_M^*,z_M^*)和(x^*,y^*,z_M^*)是问题 D_0 的可行解，则(x_M^*,y_M^*,z_M^*)是问题 D_0 的最优解，且 $y^*=y_M^*$ 和 $z^*=z_M^*$。

证明：根据定理的已知条件，我们有：

$$F(x_M^*,y_M^*,z_M^*;M,\rho)=(\sum_{i=1}^n y_{Mi}^*-M)^2\leqslant$$

$$F(x^*,y^*,z_M^*;M,\rho)=(\sum_{i=1}^n y_i^*-M)^2$$

因为 $M<0$，$\sum_{i=1}^n y_{Mi}^*\geqslant0$，$\sum_{i=1}^n y_i^*\geqslant0$，我们有 $\sum_{i=1}^n y_{Mi}^*\leqslant\sum_{i=1}^n y_i^*$，由$(x^*,y^*,z^*)$是问题 D_0 的最优解，可得$\sum_{i=1}^n y_{Mi}^*\geqslant\sum_{i=1}^n y_i^*$，因此$(x_M^*,y_M^*)$是问题 D_0 的最优解。

定理 3 给定 $M<0$，设(x_M^*,y_M^*)是问题 $D_0(M,\rho)_{z_M^*}$ 的最优解，z_M^* 是问题 $D_0(M,\rho)_{(x_M^*,y_M^*)}$ 的最优解，如果(x_M^*,y_M^*,z_M^*)是问题 D_0 的可行解，则(x_M^*,y_M^*,z_M^*)是问题 D_0 的局部最优解。

证明：设(x^*,y^*,z_M^*)是问题 D_0 中固定 z_M^* 的最优解，根据定理的已知条件，我们有：

$$F(x_M^*,y_M^*,z_M^*;M,\rho)=(\sum_{i=1}^n y_{Mi}^*-M)^2\leqslant$$

$$F(x^*,y^*,z_M^*;M,\rho)=(\sum_{i=1}^n y_i^*-M)^2$$

因为 $M<0$，$\sum_{i=1}^n y_{Mi}^*\geqslant0$，$\sum_{i=1}^n y_i^*\geqslant0$，所以有 $\sum_{i=1}^n y_{Mi}^*\leqslant\sum_{i=1}^n y_i^*$。反过来，显然$\sum_{i=1}^n y_{Mi}^*\geqslant\sum_{i=1}^n y_i^*$，因此，$(x_M^*,y_M^*)$是问题 D_0 中固定 z_M^* 的最优解，另一方面，显然 z_M^* 是问题 D_0 中固定(x_M^*,y_M^*)

的最优解，即(x_M^*,y_M^*,z_M^*)是问题 D_0 的局部最优解。

下面给出本文提出的算法，如算法 1 所示。

算法 1

Step1　设(x^1,y^1,z^1)，$M_1<0$，$N>1$，$\rho_1>0$ 和 $k=1$。

Step2　以(x^k,y^k,z^k)为起始点计算优化问题：

$$z_i^{k+1}=\frac{1-y_i^k}{1+(x_i^k)^2+(y_i^k)^2},i=1,2,\cdots,n$$

$$(x^{k+1},y^{k+1})=\mathop{\arg\min}_{(x,y)\in R^n\times R^n} F(x,y;z^{k+1};M_k,\rho_k)$$

Step3　若$(x^{k+1},y^{k+1},z^{k+1})$是问题 D_0 的可行解，则$(x^{k+1},y^{k+1},z^{k+1})$是问题 D_0 的局部最优解。否则，若 $f(x^{k+1},y^{k+1},z^{k+1})\leqslant M_k$，令 $M_{k+1}=NM_k,\rho_{k+1}=N\rho_k,k=k+1$，转 Step2；若 $f(x^{k+1},y^{k+1},z^{k+1})>M_k$，则令 $M_{k+1}=M_k,\rho_{k+1}=N\rho_k$。

文献[14]算法收敛的条件较多，这里不需要其他条件，证明下面算法收敛性定理：

定理 4　若问题 D_0 存在可行解，假设$\{(x^k,y^k,z^k)\}$是由算法 1 产生的序列，那么有以下结论。

(1)若$\{(x^k,y^k,z^k)\}$是有限序列，即在第 k 步终止，则(x^k,y^k,z^k)是问题 D_0 的局部最优解。

(2)若$\{(x^k,y^k,z^k)\}$是无限序列，假设存在某个 k' 使得对所有 $k\geqslant k'$ 都有 $f(x^{k+1},y^{k+1},z^{k+1})>M_{k'}$，如果序列$\{F(x^{k+1},y^{k+1};z^{k+1};M_k,\rho_k)\}$有界，则$\{(x^k,y^k,z^k)\}$的每个收敛聚点$(x^*,y^*,z^*)$是问题 D_0 的局部最优解，并且存在不全为 0 的常数 $\alpha_1,\alpha_2,\cdots,\alpha_m,\alpha_{m+1}$ 使得：

$$\nabla f(x^*,y^*;z^*)+\sum_{j=1}^m\alpha_j\nabla g_j(x^*)+\alpha_{m+1}\nabla h(x^*,y^*;z^*)=0 \tag{5}$$

否则，问题 D_0 无最优解。

证明：对于结论(1)，根据定理 2 可知结论成立。

对于结论(2)，假设存在某个 k' 使得对所有 $k\geqslant k'$ 都有 $f(x^{k+1},y^{k+1},z^{k+1})>M_k$，根据算法 1 有 $M_{k'}=M_k$，所以有 $k\geqslant k'$。

因序列$\{F(x^{k+1},y^{k+1};z^{k+1};M_k,\rho_k)\}$有界，存在常数 $L>0$ 使得：

$$(\sum_{i=1}^n y_i^{k+1}-M_{k'})^2\leqslant F(x^{k+1},y^{k+1};y^{k+1};M_k,\rho_k)\leqslant L$$

所以有：

$$k\geqslant k' \tag{6}$$

则有 $0<\sum_{i=1}^n y_i^{k+1}-M_{k'}<\sqrt{L}$(当 $k>k'$ 时)，因此，可得序列$\{\sum_{i=1}^n y_i^{k+1}\}$是有界的，由式(3)和式(6)得：

$$(\sum_{i=1}^n y_i^{k+1}-M_{k'})^2+\rho C(x^{k+1},y^{k+1},z^{k+1})<L$$

即

$$0\leqslant C(x^{k+1},y^{k+1},z^{k+1})<\frac{1}{\rho_k}(L-(\sum_{i=1}^n y_i^{k+1}-M_{k'})^2)$$

由上式得知序列$\{(x^{k+1},y^{k+1},z^{k+1})\}$是有界的，不妨设$(x^{k+1},y^{k+1},z^{k+1})\to(x^*,y^*,z^*)$(当 $k\to+\infty$)。根据上式令 $k\to+\infty$，有 $C(x^*,y^*,z^*)=0$，得(x^*,y^*,z^*)是问题 D_0 的可行解，由定理 3 得(x^*,y^*,z^*)是问题 D_0 的局部最优解。由算法 1 的 Step2 得到：

$$\nabla_{(x,y)}F(x^{k+1},y^{k+1};z^{k+1};M_k,\rho_k)=0$$

$$\nabla_z F(x^k,y^k;z^{k+1};M_k,\rho_k)=0$$

即

$$\rho_k\sum_{j=1}^m 2(\max\{g_j(x^{k+1}),0\}\nabla g_j(x^{k+1})+\max\{-g_j(x^{k+1}),0\}\nabla(-g_j(x^{k+1})))+\rho_k\nabla_x h(x^{k+1},y^{k+1};z^{k+1})=0$$

$$2\max\{f(x^{k+1},y^{k+1};z^{k+1})-M_k,0\}\nabla_y f(x^{k+1},y^{k+1};z^{k+1})+\rho_k\nabla_y h(x^{k+1},y^{k+1};z^{k+1})=0$$

$$\rho_k\nabla_z h(x^k,y^k;z^{k+1})=0$$

对于 $k>k'$，有 $2\max\{f(x^{k+1},y^{k+1};z^{k+1})-M_k,0\}>0$，令：

$$\gamma^{k+1}=1+\rho_k\sum_{j=1}^m 2(\max\{g_j(x^{k+1}),0\}+\max\{-g_j(x^{k+1}),0\})+\rho_k \tag{7}$$

有 $\gamma^{k+1}>0$，再令：

$$\beta_j^{k+1}=\frac{2\rho_k\max\{g_j(x^{k+1}),0\}}{\gamma^{k+1}}$$

$$\eta_j^{k+1}=\frac{2\rho_k\max\{-g_j(x^{k+1}),0\}}{\gamma^{k+1}},j=1,2,\cdots,m \tag{8}$$

$$\beta_{m+1}^{k+1}=\frac{\rho_k}{\gamma^{k+1}}\lambda_{k+1}=\frac{2\max\{f(x^{k+1},y^{k+1};z^{k+1})-M_k,0\}}{\gamma^{k+1}}$$

$$\eta^{k+1}=\frac{1}{\gamma^{k+1}} \tag{9}$$

有 $1=\eta_j^{k+1}+\sum_{j=1}^m(\beta_j^{k+1}+\eta_j^{k+1})+\beta_{m+1}^{k+1}$，因此，序列$\{\beta_j^{k+1}\}$，$\{\eta_j^{k+1}\}(j=1,2,\cdots,m)$，$\{\beta_{m+1}^{k+1}\}$，$\{\eta^{k+1}\}$ 收敛，当 $k\to\infty$ 时，令 $\beta_j^{k+1}\to\beta_j,\eta_j^{k+1}\to\eta_j(j=1,2,\cdots,m),\beta_{m+1}^{k+1}\to\beta_{m+1},\eta^{k+1}\to\eta$。由于$(x^{k+1},y^{k+1},z^{k+1})\to(x^*,y^*,z^*)$，有 $\lambda_{k+1}\to\lambda$，因此，由式(7)－式(9)有：

$$\sum_{j=1}^m(\beta_j^{k+1}\nabla g_j(x^{k+1})+\eta_j^{k+1}\nabla(-g_j(x^{k+1})))+\beta_{m+1}^{k+1}\nabla_x h(x^{k+1},y^{k+1};z^{k+1})=0$$

$$\lambda^{k+1}\nabla_y f(x^{k+1},y^{k+1};z^{k+1})+\beta_{m+1}^{k+1}\nabla_y h(x^{k+1},y^{k+1};z^{k+1})=0$$

$$\beta_{m+1}^{k+1}\nabla_z h(x^k,y^k;z^{k+1})=0$$

令 $k\to\infty$，有：

$$\sum_{j=1}^m(\beta_j\nabla g_j(x^*)+\eta_j\nabla(-g_j(x^*)))+\beta_{m+1}\nabla_x h(x^*,y^*;z^*)=0$$

$$\lambda\nabla_y f(x^*,y^*;z^*)+\beta_{m+1}\nabla_y h(x^*,y^*;z^*)=0$$

$$\beta_{m+1}\nabla_z h(x^*,y^*;z^*)=0$$

令 $\alpha_j=\frac{\beta_j-\eta_j}{\lambda},\alpha_{m+1}=\frac{\beta_{m+1}}{\lambda}\geqslant0,j=1,2,\cdots,m$，由上式得到：$\nabla f(x^*,y^*;z^*)+\sum_{j=1}^m\alpha_j\nabla g_j(x^*)+\alpha_{m+1}\nabla h(x^*,y^*;z^*)=0$。

最后，假设不存在某个 k' 使得对所有 $k>k'$ 都有 $f(x^{k+1},y^{k+1},z^{k+1})>M_k$，由于 $k\to+\infty,M_k\to-\infty$，则 $f(x^{k+1},y^{k+1},z^{k+1})\to-\infty$，即问题 D_0 无最优解。证毕。

定理 4 说明，通过算法 1 可以得到问题 P_0 的局部最优解或 KKT 解。

定理 5　若问题 D_0 存在可行解，假设$\{(x^k,y^k,z^k)\}$是由算法 1 产生的序列，那么有以下结论。若$\{(x^k,y^k,z^k)\}$是无限序列，假设存在某个 k' 使得对于所有 $k\geqslant k'$ 都有 $f(x^{k+1},y^{k+1},z^{k+1})>M_{k'}$，则对于 $i=1,2,\cdots,n$，有下面等式成立：

$$z_i^{k+1}=\frac{1-y_i^k}{1+(x_i^k)^2+(y_i^k)^2} \tag{10}$$

$$x^{k+1}=(A^TA+Z^{k+1})^{-1}A^Tb \tag{11}$$

$$y_i^{k+1}=\frac{1}{((z_i^{k+1})^2+1)\rho_k}[M_k+\rho_k(1-z_i^{k+1})-Y^{k+1}] \tag{12}$$

其中，$Y^{k+1}=\sum_{i=1}^n y_i^{k+1}=\sum_{i=1}^n\frac{M_k+\rho_k(1-z_i^{k+1})}{((z_i^{k+1})^2+1)}(\sum_{i=1}^n\frac{1}{(z_i^{k+1})^2+1}+\rho_k)^{-1}$，$A^TA+Z^{k+1}$是可逆的，并且有：

$$Z^{k+1}=\begin{pmatrix}(z_1^{k+1})^2 & 0 & \cdots & 0\\ 0 & (z_2^{k+1})^2 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & (z_n^{k+1})^2\end{pmatrix} \tag{13}$$

进一步,如果$(x^{k+1},y^{k+1},z^{k+1})\to(x^*,y^*,z^*)$(当 $k\to+\infty$),$M_k\to M^*$,$\rho_k\to+\infty$,则对于 $i=1,2,\cdots,n$,有下面等式成立:

$$z_i^*=\frac{1-y_i^*}{1+(x_i^*)^2+(y_i^*)^2} \tag{14}$$

$$x^*=(A^{\mathrm{T}}A+Z^*)^{-1}A^{\mathrm{T}}b \tag{15}$$

$$y_i^*=\frac{1-z_i^*}{(z_i^*)^2+1} \tag{16}$$

其中,$\sum_{i=1}^n y_i^*=\sum_{i=1}^n\frac{(1-z_i^*)}{((z_i^*)^2+1)}$,并且有:

$$Z^*=\begin{pmatrix}(z_1^*)^2 & 0 & \cdots & 0\\ 0 & (z_2^*)^2 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & (z_n^*)^2\end{pmatrix} \tag{17}$$

其中,(x^*,y^*,z^*)是问题 D_0 的局部最优解。

证明:根据算法 1 的 Step2,对于 $i=1,2,\cdots,n$,对

$$F(x,y;z^{k+1};M_k,\rho_k)=(\sum_{i=1}^n y_i-M_k)^2+\rho_k\sum_{i=1}^n[(x_i^2+y_i^2)(z_i^{k+1})^2+(y_i+z_i^{k+1}-1)^2]+\rho_k\sum_{j=1}^m(A_jx-b_j)^2$$

关于(x,y)求导:$\frac{\partial F}{\partial x_i}=0$,$\frac{\partial F}{\partial y_i}=0$,得到:

$$\rho\sum_{i=1}^n 2x_i^{k+1}(z_i^{k+1})^2+\rho\sum_{j=1}^m 2(A_jx^{k+1}-b_j)a_{ji}=0 \tag{18}$$

$$2(\sum_{i=1}^n y_i^{k+1}-M)+\rho_k 2y_i^{k+1}(z_i^{k+1})^2+\rho_k 2(y_i^{k+1}+z_i^{k+1}-1)=0 \tag{19}$$

由式(18)得到$\sum_{j=1}^m A_j^{\mathrm{T}}(A_jx^{k+1}-b_j)+Z^{k+1}x^{k+1}=0$,即$(A^{\mathrm{T}}A+Z^{k+1})x^{k+1}=A^{\mathrm{T}}b$(得式(11)),其中 Z^{k+1} 按式(13)定义。由式(19)得:

$$\frac{1}{(z_i^{k+1})^2+1}(\sum_{i=1}^n y_i^{k+1}-M)+\rho_k y_i^{k+1}+\rho_k\frac{1}{(z_i^{k+1})^2+1}(z_i^{k+1}-1)=0,i=1,2,\cdots,n \tag{20}$$

式(20)的 n 个等式相加得:

$$(\sum_{i=1}^n\frac{1}{(z_i^{k+1})^2+1}+\rho_k)\sum_{i=1}^n y_i^{k+1}=\sum_{i=1}^n\frac{M_k+(1-z_i^{k+1})\rho_k}{(z_i^{k+1})^2+1}$$

即得:

$$Y^{k+1}=\sum_{i=1}^n y_i^{k+1}=\sum_i^n\frac{M_k+\rho_k(1-z_i^{k+1})}{((z_i^{k+1})^2+1)}(\sum_i^n\frac{1}{(z_i^{k+1})^2+1}+\rho_k)^{-1} \tag{21}$$

将式(21)代入式(20)得到式(12)。由式(10)知 $z_i^{k+1}=\frac{1-y_i^k}{1+(x_i^k)^2+(y_i^k)^2}<1$,根据式(21)得:

$$\begin{aligned}Y^{k+1}&=\sum_{i=1}^n y_i^{k+1}\\&=\sum_i^n\frac{M_k/\rho_k+(1-z_i^{k+1})}{((z_i^{k+1})^2+1)}(\frac{1}{\rho_k}\sum_i^n\frac{1}{(z_i^{k+1})^2+1}+1)^{-1}\\&<\sum_i^n\frac{M_k/\rho_k+(1-z_i^{k+1})}{((z_i^{k+1})^2+1)}<\sum_i^n\frac{(1-z_i^{k+1})}{((z_i^{k+1})^2+1)}<n\end{aligned}$$

则序列$\{F(x^{k+1},y^{k+1};z^{k+1};M_k,\rho_k)\}$有界。另一方面,因为 $M_k<0$,由式(21)得:

$$\begin{aligned}Y^{k+1}&=\sum_{i=1}^n y_i^{k+1}\\&=\sum_i^n\frac{M_k/\rho_k+(1-z_i^{k+1})}{((z_i^{k+1})^2+1)}(\frac{1}{\rho_k}\sum_i^n\frac{1}{(z_i^{k+1})^2+1}+1)^{-1}\\&>\sum_i^n\frac{M_k}{((z_i^{k+1})^2+1)}(\sum_i^n\frac{1}{(z_i^{k+1})^2+1}+1)^{-1}>M_k\end{aligned}$$

由定理 4 得式(14)-式(17)成立,且(x^*,y^*,z^*)是问题 D_0 的局部最优解。

根据定理 5,若 x^* 是(P_0)的最优解,$x_i^*=0$,则令 $y_i^*=0$,否则 $y_i^*=1$。然后令 $z_i^*=1-y_i^*$,$i=1,2,\cdots,n$,那么 $A^{\mathrm{T}}Ax^*+Z^*x^*=A^{\mathrm{T}}b$ 成立,若$(A^{\mathrm{T}}A+Z^*)^{-1}$存在,则有式(15)成立。

据此,本文构建一个交替随机搜索算法。

算法 2

Step1 设 $k=1$ 和 $s_1=n$,给定误差值 ε;

Step2 随机生成 $y^k=(y_1^k,y_2^k,\cdots,y_n^k)^{\mathrm{T}}$,其中 y_i^k 随机取 1 的总数不超过 s_k,其余 $y_0^k=0$。然后令 $z_i^k=1-y_i^k$,$i=1,2,\cdots,n$,计算 $x^k=(A^{\mathrm{T}}A+Z^k)^{-1}A^{\mathrm{T}}b$。转 Step3。

Step3 如果$|x_i^k|<\varepsilon$,则令 $y_i^{k+1}=0$,否则 $y_i^{k+1}=1$,然后令 $z_i^k=1-y_i^k$,$i=1,2,\cdots,n$,计算 $x^{k+1}=(A^{\mathrm{T}}A+Z^{k+1})^{-1}A^{\mathrm{T}}b$,$s_{k+1}=\sum_{i=1}^n y_i^{k+1}$。令 $k:=k+1$,若 s_k 不变化则终止,否则转 Step2。

通过算法 2 的 Step2 得到一个单调递减序列$\{s_k\}$,下面给出一个算例说明算法 1 的使用过程。

例 1 给出一个简单 0 模问题中算法 2 的使用过程。

$\min\ \|x\|_0$

s. t. $2x_1+x_2+x_3=1,x_1+x_2+2x_3=1$。

显然存在一个最优解$(x_1,x_2,x_3)=(0,1,0)^{\mathrm{T}}$,它的等价表示为:

$\min\ y_1+y_2+y_3$

s. t. $2x_1+x_2+x_3=1,x_1+x_2+2x_3=1$,

$x_iz_i=0,y_iz_i=0,z_i+y_i-1=0,i=1,2,3$

显然存在一个最优解:$(x_1,x_2,x_3)=(0,1,0)^{\mathrm{T}}$,若取$(y_1,y_2,y_3)=(0,1,0)^{\mathrm{T}}$,$(z_1,z_2,z_3)=(1,0,1)^{\mathrm{T}}$,其中 $A=\begin{pmatrix}2&1&1\\1&1&2\end{pmatrix}$,$b=\begin{pmatrix}1\\1\end{pmatrix}$,由式(15)得:

$$\begin{aligned}x^*&=(A^{\mathrm{T}}A+Z^*)^{-1}A^{\mathrm{T}}b\\&=\left(\begin{pmatrix}2&1&1\\1&1&2\end{pmatrix}^{\mathrm{T}}\begin{pmatrix}2&1&1\\1&1&2\end{pmatrix}+\begin{pmatrix}1&0&0\\0&0&0\\0&0&1\end{pmatrix}\right)^{-1}\begin{pmatrix}2&1&1\\1&1&2\end{pmatrix}^{\mathrm{T}}\begin{pmatrix}1\\1\end{pmatrix}\\&=\begin{pmatrix}3/4&-3/2&1/4\\-3/2&5&-3/2\\1/4&-3/2&3/4\end{pmatrix}\begin{pmatrix}2&1&1\\1&1&2\end{pmatrix}^{\mathrm{T}}\begin{pmatrix}1\\1\end{pmatrix}\\&=\begin{pmatrix}0\\1\\0\end{pmatrix}\end{aligned}$$

例 2 设问题 $\min\ \|x\|_0$ s. t. $Ax=1$,其中 A 是 $m\times n$ 阶随机正态分布矩阵,$n=2m$ 和第 m 例所有的元素用 1 取代,

即$(a_{1m},a_{2m},\cdots,a_{mm})^{\mathrm{T}}=(1,1,\cdots,1)^{\mathrm{T}}$，显然该问题的最优解 x^* 为第 m 维等于 1，其余维等于 0，即 $\|x^*\|_0=1$。针对算法 2，用 MATLAB2014a 计算得到表 1，设置误差 $\varepsilon=0.000001$。针对下面 7 个维度的问题，每个维度随机计算 100 个模型，搜索到最优解的成功率均为 100%，平均迭代次数为 4－5。例如对于 800 维的 100 个模型的计算结果，最大迭代次数不超过 20(见图 1)，平均所用时间不超过 0.87 s(见图 2)。

表 1　$n=2m$ 的 100 个模型的迭代结果比较

n	最大迭代次数	平均迭代次数	最大迭代时间/s	平均迭代时间/s
10	15	4.54	0.015	0.002
100	13	3.56	0.013	0.096
200	18	4.16	0.194	0.046
400	12	3.49	0.414	0.115
800	13	3.88	1.882	0.531
1600	19	4.31	12.011	2.852
3200	32	4.84	118.757	16.704

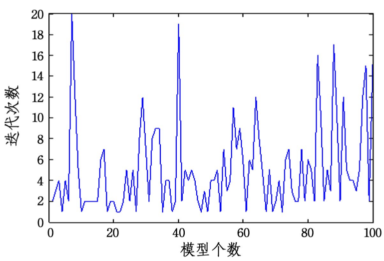

图 1　当 $n=800$ 维时 100 个模型收敛到最优解的迭代次数

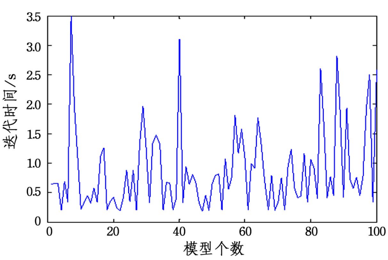

图 2　当 $n=800$ 维时 100 个模型收敛到最优解时的计算时间

算法 2 主要对如例 2 中的特定压缩感知问题具有较好的收敛性，说明当问题规模不大时，足够多的遍历可以获得原问题的最优解。因此，对于其他实际问题需要改造算法 2 中的 Step2，比如按遗传算法的交叉变异等，这些有待于进一步研究。

结束语　本文提出了压缩感知问题的一种双凸优化问题的等价表示，定义了一个对应的目标罚函数算法，证明了该算法的收敛性，进一步推导了压缩感知最优解的解析表达式，提出了一个交替随机搜索算法。该算法对于一些特殊压缩感知问题有效。

本文得到的压缩感知问题最优解的解析形式有助于进一步设计新的算法，该文的思路为研究压缩感知问题开辟了一个新的方向，为研究稀疏优化问题提供了新的思路。

参 考 文 献

[1] CANDÈS E,TAO T. Near optimal signal recovery from random projections:Universal encoding strategies [J]. IEEE Trans. Info. Theory,2006,52(12):5406-5425.

[2] CANDÈS E J,WAKIN M B,BOYD S P. Enhancing Sparsity by Reweighted l_1 Minimization[J]. J. Fourier Anal. Appl.,2008,14:877-905.

[3] CHARTRAND R,YIN W. Iteratively reweighted algorithms for compressive sensing[C]// Proc. IEEE Int. Conf. Acoust,Speech,Signal Process. 2008:3869-3872.

[4] MOHIMANI H,BABIE-ZADEH M,JUTTEN C. A fast approach for overcomplete sparse decomposition based on smoothed l_0-norm [J]. IEEE Trans. Signal Process.,2009,57(1):289-301.

[5] FOUCART S,RAUHUT H. A Mathematical Introduction to Compressive Sensing[M]. Springer,New York,2013.

[6] PANT J K,LU W S,ANTONIOU A. New Improved Algorithms for Compressive Sensing Based on l_p Norm[J]. IEEE Transactions on Circuits and Systems－II:Express Briefs,2014,61(3):198-202.

[7] WANG Y,WANG J J,XU Z B. Restricted p-isometry properties of nonconvex block-sparse compressed sensing[J]. Signal Processing,2014,104:188-196.

[8] ZHU Y,WU J,YU G H. A fast proximal point algorithm for l_1-minimization problem mincompressed sensing [J]. Applied Mathematics and Computation,2015,270:777-784.

[9] 文婷婷，马兆楠，裴炳，基于拟牛顿法的压缩感知重构零范数平滑算法[J]. 计算机应用，2015,35(S2):17-19,23.

[10] 杜卓明，李洪安，康宝生，等. 二阶收敛的光滑正则化压缩感知信号重构方法[J]. 中国图象图形学报，2016,21(4):490-498.

[11] MENG Z Q,DANG C Y,JIANG M,et al. Exactness and algorithm of an objective penalty function[J]. Journal Globel Optimization,2013,56:691-711.

[12] GORSKI J,PFEUFFER F,KLAMROTH K. Biconvex sets and optimization with biconvex functions:a survey and extensions [J]. Math. Meth. Oper. Res.,2007,66:373-407.

[13] 孟志青，徐蕾艳，蒋敏，等. 压缩感知优化问题的等价表示及其目标罚函数方法[J]. 计算机科学，2017,44(S1):97-99.