

基于代理函数优化的稀疏性字典学习

高 磊, 陈曾平

(国防科技大学 ATR 重点实验室, 湖南长沙 410073)

摘 要: 稀疏性字典学习是指对在某个已知的基字典上具有稀疏表示的字典的学习. 论文利用块松弛思想, 将稀疏性字典学习问题转化为字典和系数的分别优化问题, 利用代理函数优化方法分别对固定字典和固定系数情况下的目标函数进行优化处理, 得到固定字典情况下的系数更新算法和固定系数情况下的字典更新算法, 进而得到稀疏性字典学习算法. 理论分析说明了本文算法的收敛性. 仿真对比表明了本文算法在收敛性和运算效率方面均优于稀疏性 K-SVD 算法.

关键词: 稀疏表示; 稀疏性字典; 块松弛; 代理函数; K-SVD

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 0372-2112 (2011) 12-2910-04

Sparse Dictionary Learning Based on Optimization of Surrogate Function

GAO Lei, CHEN Zeng-ping

(ATR Key Lab, National University of Defense Technology, Changsha, Hunan 410073, China)

Abstract: Sparse dictionary learning means that learning for a dictionary which has sparse representation on a known base dictionary. In the paper, with block-relaxation, the sparse dictionary learning can be translated into respective optimization of dictionary and coefficients. It means that the target function can be optimized respectively with fixed dictionary or fixed coefficients by optimization method of surrogate function. Through above process, the update algorithm of coefficients with fixed dictionary and update algorithm of dictionary with fixed coefficients can be obtained. Then the sparse dictionary learning algorithm is obtained. The convergence of the algorithm is illuminated theoretically. Comparison in simulation indicates that the algorithm put forward in this paper is superior to sparse K-SVD algorithm in convergence and operating efficiency.

Key words: sparse representation; sparse dictionary; block-relaxation; surrogate function; K-SVD

1 引言

字典学习是近年来稀疏表示领域中的热点问题^[1], 所提出的方法主要有 K-SVD 方法^[2]、基于代理函数优化的方法^[3]、递归最小二乘方法^[4]、迭代子空间法^[5]以及贝叶斯正交成份分析法^[6].

文献[7]在稀疏性字典模型下考虑字典学习问题. “稀疏性”是指字典 D 满足 $D = \Phi A$, 其中 $\Phi \in R^{d \times l}$ 为预先选取的基字典, $A \in R^{l \times k}$ 为字典 D 在 Φ 上的稀疏表示矩阵(以下称:字典稀疏矩阵), 其每一列为稀疏序列.

针对稀疏性字典的学习问题, 文献[7]给出了稀疏性 K-SVD 算法. 稀疏性 K-SVD 算法沿用了文献[2]中的 K-SVD 算法的原理. 文献[3]的作者观察到 K-SVD 算法在最佳解的邻域内的稀疏逼近误差未必会在每一步迭代中减小, 这使得 K-SVD 算法最终得到的解未必是最佳的, 由于稀疏性 K-SVD 算法与 K-SVD 的相似性, 稀疏性 K-SVD 算法得到的稀疏性字典也未必是最佳的. 鉴于此, 借鉴文献[3]的整体性优化的思想, 本文利用块松弛和代理函数优化的方法, 发展了一种新的基于代理函数优化的

稀疏性字典学习算法(Sparse Dictionary Learning Algorithm based on Optimization of Surrogate Function, SDLA-OSF). SDLA-OSF 将字典和使用字典的系数分别作为一个整体进行更新, 更新的过程中保持了字典的稀疏性结构; 仿真表明, SDLA-OSF 收敛性能要优于稀疏性 K-SVD 算法.

2 问题描述

设字典学习的目标为稀疏性字典: $D = \Phi A$, 其中 Φ 为已知基字典, A 为字典稀疏矩阵. 设训练样本矩阵为: $Y \in R^{d \times N}$, Y 在 D 上的稀疏表示系数矩阵为 $X \in R^{k \times N}$ (以下称为:系数矩阵), 则本文解决的主要问题是:

$$(DL_1) \min_{A, X} \phi(A, X), \text{ s. t.}$$

$$\phi(A, X) = \|Y - \Phi AX\|_F^2 + \lambda_X J(X) + \lambda_A J(A) \quad (1)$$

其中 $\|\cdot\|_F$ 表示矩阵的 Frobenius 范数(F 范数), λ_X, λ_A 分别为加权系数, 为正值, $J(\cdot)$ 为矩阵稀疏度的 l^1 范数度量函数^[8], 其定义如下:

$$J(X) = \sum_{i \in \text{Row}_X} \sum_{j \in \text{Col}_X} |x_{i,j}| \quad (2)$$

$\text{Row}_X, \text{Col}_X$ 分别为矩阵 X 的行序号和列序号的集合.

该问题能够使用凸优化方法求解. 基于块松弛思想的优化方法通常分为两步交替进行, 以收敛到最优解: (1) 固定 \mathbf{A} , 更新 \mathbf{X} ; (2) 固定 \mathbf{X} , 更新 \mathbf{A} . 然而在每一步中, $\phi(\mathbf{A}, \mathbf{X})$ 中 \mathbf{A}, \mathbf{X} 是相互耦合的, 难以直接求解, 因此本文考虑用代理函数优化方法^[3]求解上述更新问题.

3 基于代理函数优化的稀疏性字典学习算法

3.1 代理函数优化的基本原理^[3]

代理函数优化方法的主要思想将参数互相耦合难以优化的目标函数 ϕ 用一个参数相互解耦的易于优化的代理函数 ψ 代替, ψ 必须满足:

$$\phi(\omega) \leq \psi(\omega, \xi), \forall \omega, \xi \in \mathcal{T}; \phi(\omega) = \psi(\omega, \omega), \forall \omega \in \mathcal{T} \quad (3)$$

其中 \mathcal{T} 是参数集合. ξ 为与 ω 解耦的辅助参数.

代理函数优化的基本流程为:

(1) 给定参数 ξ 的初始值, 使其等于 ω 的当前值 ω_{cur} , 通过优化的方法寻找 ω 的最优更新:

$$\omega_{\text{new}} = \arg \min_{\omega \in \mathcal{T}} \psi(\omega, \xi), \xi = \omega_{\text{cur}} \quad (4)$$

(2) 更新 $\xi = \omega_{\text{new}}$, 重复迭代直到收敛.

交替迭代满足 $\psi(\omega_{\text{new}}, \omega_{\text{cur}}) \leq \psi(\omega_{\text{cur}}, \omega_{\text{cur}})$, 由式(3)有: $\phi(\omega_{\text{new}}) \leq \psi(\omega_{\text{new}}, \omega_{\text{cur}})$, $\psi(\omega_{\text{cur}}, \omega_{\text{cur}}) = \phi(\omega_{\text{cur}})$, 故 $\phi(\omega_{\text{new}}) \leq \phi(\omega_{\text{cur}})$, 即 $\phi(\omega)$ 在迭代过程中是不增的.

3.2 稀疏性字典学习的代理函数构造

针对一般的字典的学习, 文献[3]构造了代理函数并给出了系数矩阵和字典的更新算法. 本文将其构造方法推广到稀疏性字典学习中.

根据式(1)的优化问题, 构造两个代理函数:

(1) \mathbf{A} 固定时, 代理函数 $\psi_X(\mathbf{X}, \mathbf{X}^{[n-1]})$ 为:

$$\begin{aligned} \psi_X(\mathbf{X}, \mathbf{X}^{[n-1]}) = & \|\mathbf{Y} - \Phi\mathbf{A}\mathbf{X}\|_F^2 + c_D \|\mathbf{X} - \mathbf{X}^{[n-1]}\|_F^2 \\ & - \|\Phi\mathbf{A}\mathbf{X} - \Phi\mathbf{A}\mathbf{X}^{[n-1]}\|_F^2 \\ & + \lambda_X J(\mathbf{X}) + \lambda_A J(\mathbf{A}) \end{aligned} \quad (5)$$

其中 $\mathbf{X}^{[n-1]}$ 为上一步迭代得到的系数矩阵. $c_D \geq \|(\Phi\mathbf{A})^T(\Phi\mathbf{A})\|$ 为常数 ($\|\cdot\|$ 表示矩阵的谱范数), λ_X, λ_A 分别为 \mathbf{X} 和 \mathbf{A} 稀疏度函数的加权系数.

(2) \mathbf{X} 固定时, 代理函数 $\psi_A(\mathbf{A}, \mathbf{A}^{[n-1]})$ 为:

$$\begin{aligned} \psi_A(\mathbf{A}, \mathbf{A}^{[n-1]}) = & \|\mathbf{Y} - \Phi\mathbf{A}\mathbf{X}\|_F^2 + c_{\Phi\mathbf{X}} \|\mathbf{A} - \mathbf{A}^{[n-1]}\|_F^2 \\ & - \|\Phi\mathbf{A}\mathbf{X} - \Phi\mathbf{A}^{[n-1]}\mathbf{X}\|_F^2 \\ & + \lambda_X J(\mathbf{X}) + \lambda_A J(\mathbf{A}) \end{aligned} \quad (6)$$

其中 $\mathbf{A}^{[n-1]}$ 为上一步迭代得到的字典稀疏矩阵, $c_{\Phi\mathbf{X}} = c_X c_{\Phi}$, $c_{\Phi} \geq \|\Phi^T \Phi\|$, $c_X \geq \|\mathbf{X}^T \mathbf{X}\|$, 其余的符号含义同式(5).

计算可知, 代理函数对 \mathbf{A} 或 \mathbf{X} 的梯度实现了 \mathbf{A} 和 \mathbf{X} 的解耦. 根据定义, $\psi_X(\mathbf{X}, \mathbf{X}^{[n-1]})$ 和 $\psi_A(\mathbf{A}, \mathbf{A}^{[n-1]})$ 必须满足式(3), 则下述两个不等式须成立:

$$c_D \|\mathbf{X} - \mathbf{X}^{[n-1]}\|_F^2 - \|\Phi\mathbf{A}\mathbf{X} - \Phi\mathbf{A}\mathbf{X}^{[n-1]}\|_F^2 \geq 0 \quad (7)$$

$$c_{\Phi\mathbf{X}} \|\mathbf{A} - \mathbf{A}^{[n-1]}\|_F^2 - \|\Phi\mathbf{A}\mathbf{X} - \Phi\mathbf{A}^{[n-1]}\mathbf{X}\|_F^2 \geq 0 \quad (8)$$

文献[3]利用泰勒级数推导了与式(7)、(8)类似的不等式, 读者可结合文献[3]进行推导.

3.3 基于代理函数的稀疏性字典学习

稀疏性字典学习分为两个关键性步骤:

(1) 字典固定, 求解系数矩阵的更新算法;

(2) 系数矩阵固定, 求解字典的更新算法.

下面根据凸优化理论, 给出学习算法.

3.3.1 固定字典, 更新系数矩阵

利用凸优化理论最小化式(5). 由凸函数的性质可知, 代理函数相对 \mathbf{X} 的极小值点意味着代理函数相对于 \mathbf{X} 的梯度为零^[9]. 计算 $\psi_X(\mathbf{X}, \mathbf{X}^{[n-1]})$ 对 \mathbf{X} 的梯度, 可得下式:

$$\begin{aligned} \nabla_X \psi_X(\mathbf{X}, \mathbf{X}^{[n-1]}) & = 2c_D \mathbf{X} - 2(\Phi\mathbf{A})^T \mathbf{Y} - 2c_D \mathbf{X}^{[n-1]} + 2(\Phi\mathbf{A})^T (\Phi\mathbf{A}) \mathbf{X}^{[n-1]} \\ & \quad + \lambda_X \nabla_X J(\mathbf{X}) \end{aligned} \quad (9)$$

令 $\mathbf{B} = 1/c_D ((\Phi\mathbf{A})^T \mathbf{Y} + c_D \mathbf{X}^{[n-1]} - (\Phi\mathbf{A})^T (\Phi\mathbf{A}) \mathbf{X}^{[n-1]})$, 将 \mathbf{B} 记为 $\mathbf{B} = [b_{i,j}]$, 由极小点条件可知: $\nabla_X \psi_X(\mathbf{X}, \mathbf{X}^{[n-1]}) = 0$, 设系数矩阵更新结果记为 $\mathbf{X}^{[n]} = [x_{i,j}^{[n]}]$, 由文献[10], 可得系数矩阵的更新公式如下:

$$[x_{i,j}^{[n]}] = \begin{cases} b_{i,j} - \frac{\lambda_X \text{sign}(b_{i,j})}{2c_D} \frac{\lambda_X}{2c_D} < |b_{i,j}|, & i \in \text{Row}_B \\ & j \in \text{Col}_B \\ 0, & \text{其他} \end{cases} \quad (10)$$

其中 $\text{Row}_B, \text{Col}_B$ 为矩阵 \mathbf{B} 的行序号和列序号的集合. 上式可简记为: $\mathbf{X}^{[n]} = S_{\lambda_X, c_D}(\mathbf{B})$.

在上述更新公式基础上, 对于固定字典情况, 更新系数矩阵可按下述算法进行:

算法 1 系数矩阵更新算法

- 步骤 1 初始化: $\mathbf{A} = \mathbf{A}_{t-1}, \mathbf{X}^{[0]} = \mathbf{X}_{t-1}, c_D \geq \|(\Phi\mathbf{A}_{t-1})^T(\Phi\mathbf{A}_{t-1})\|$, 迭代次数 K_X , 迭代计数器 $n = 0$
- 步骤 2 $n = n + 1$
- 步骤 3 如果 $n \leq K_X$: 计算 $\mathbf{B} = 1/c_D ((\Phi\mathbf{A})^T \mathbf{Y} + c_D \mathbf{X}^{[n-1]} - (\Phi\mathbf{A})^T (\Phi\mathbf{A}) \mathbf{X}^{[n-1]})$, $\mathbf{X}^{[n]} = S_{\lambda_X, c_D}(\mathbf{B})$, 转步骤 2; 如果 $n > K_X$: 转入步骤 4;
- 步骤 4 输出系数矩阵迭代结果: $\mathbf{X}_t = \mathbf{X}^{[n-1]}$, 结束算法.

3.3.2 固定系数矩阵, 更新字典

用凸优化理论最小化式(6), 类似式(10)的推导, 令 $\mathbf{G} = 1/(c_{\Phi\mathbf{X}}) (\Phi^T \mathbf{Y} \mathbf{X}^T + c_{\Phi\mathbf{X}} \mathbf{A}^{[n-1]} - \Phi^T \Phi \mathbf{A}^{[n-1]} \mathbf{X} \mathbf{X}^T)$, 将 \mathbf{G} 记为 $\mathbf{G} = [g_{i,j}]$, 设得到的字典稀疏矩阵更新记为 $\mathbf{A}^{[n]} = [a_{i,j}^{[n]}]$, 则字典稀疏矩阵更新公式如下:

$$a_{i,j}^{[n]} = \begin{cases} g_{i,j} - \frac{\lambda_A \text{sign}(g_{i,j})}{2c_{\Phi\mathbf{X}}} \frac{\lambda_A}{2c_{\Phi\mathbf{X}}} < |g_{i,j}|, & i \in \text{Row}_G; j \in \text{Col}_G \\ 0, & \text{其他} \end{cases} \quad (11)$$

其中 $\text{Row}_G, \text{Col}_G$ 为矩阵 G 的行序号和列序号的集合. 上式可简记为: $A^{[n]} = S_{\lambda_A, c_{\Phi X}}(G)$.

下面给出字典稀疏矩阵的更新算法:

算法 2 字典稀疏矩阵更新算法

- 步骤 1 初始化: $A^{[0]} = A_{i-1}, X = X_i, c_X \geq \|X^T X_i\|, c_{\Phi} \geq \|\Phi^T \Phi\|, c_{\Phi X} = c_X c_{\Phi}$, 迭代次数 K_A , 迭代计数器 $n = 0$
- 步骤 2 $n = n + 1$
- 步骤 3 如果 $n \leq K_A$: 计算 $G = 1/(c_{\Phi X})(\Phi^T Y X^T + c_{\Phi X} A^{[n-1]} - \Phi^T \Phi A^{[n-1]} X X^T)$, $A^{[n]} = S_{\lambda_A, c_{\Phi X}}(G)$, 转入步骤 2;
如果 $n > K_A$: 转入步骤 4;
- 步骤 4 输出字典稀疏矩阵迭代结果: $A_i = A^{[n-1]}$, 结束算法.

3.3.3 稀疏性字典学习算法

在上述两个子算法的基础上, 最终的稀疏性字典学习算法可描述如下:

算法 3 稀疏性字典学习算法

- 步骤 1 初始化字典和系数矩阵: A_0, X_0 和总的迭代次数 K_T
- 步骤 2 $t = t + 1$
- 步骤 3 如果 $t \leq K_T$: 在 X_{t-1}, A_{t-1} 基础上, 由算法 1 更新系数矩阵得到 X_t ; 在 X_t, A_{t-1} 基础上, 由算法 2 更新字典矩阵, 得到 A_t ; 转入步骤 2;
如果 $t > K_T$: 转入步骤 4;
- 步骤 4 给出输出: 字典稀疏矩阵 A_{t-1} , 系数矩阵 X_{t-1} , 结束算法.

3.4 稀疏性字典学习算法收敛性

SDLA-OSF 算法的收敛性能分析主要包括算法的稳定性和收敛性.

3.4.1 算法稳定性分析

首先分析算法的稳定性. $\psi_X(X, X^{[n-1]})$ 相对于 X 的 Hessian 矩阵为:

$$H[\psi_X(X, X^{[n-1]})] = \frac{\partial(\text{vec}(2c_D X))}{\partial(\text{vec}(X))^T} = 2c_D E_{NK \times NK} \quad (12)$$

上式中 $E_{NK \times NK}$ 表示的是 $NK \times NK$ 维单位矩阵, 这意味着 $\psi_X(X, X^{[n-1]})$ 的 Hessian 矩阵是正定矩阵. 由 Lyapunov 第二定理^[11], 算法 1 迭代过程稳定, 类似的, 可以证明算法 2 迭代过程稳定.

3.4.2 算法收敛性分析

构造组合优化函数:

$$\begin{aligned} \phi(A, X, A^+, X^+) &= \phi(A, X) + c_D \|X - X^+\|_F^2 \\ &\quad + c_{\Phi X} \|A - A^+\|_F^2 \\ &\quad - \|\Phi A X - \Phi A^+ X^+\|_F^2 \end{aligned} \quad (13)$$

其中 $\phi(A, X) = \|Y - \Phi A X\|_F^2 + \lambda_X J(X) + \lambda_A J(A)$. 若固定 $(A, A^+ |_{A^+ = A})$, 上式等价于式(5), 若固定 $(X, X^+ |_{X^+ = X})$, 上式等价于式(6).

式(13)的收敛性证明可参照文献[3]的思路, 将迭代过程用算子来描述, 应用闭图像定理实现证明, 限于篇幅, 本文不再给出证明详细过程.

4 仿真分析

将本文的 SDLA-OSF 与文献[7]中的稀疏性 K-SVD 算法进行仿真对比. 稀疏性 K-SVD 算法解决的问题是在限定 X, A 的列向量稀疏度情况(对 X 的任一列 $\|x_i\|_0 \leq T_X, A$ 的任一列 $\|a_j\|_0 \leq T_A, \|\cdot\|_0$ 表示非零值的个数)下, 最小化的目标函数是稀疏表示误差 $\|Y - \Phi A X\|_F^2$.

在引言中提到, 稀疏性 K-SVD 算法在迭代过程中稀疏表示误差随迭代过程中可能发生振荡. SDLA-OSF 的代理函数在迭代过程中是收敛的, 下面通过仿真进行验证. 仿真软件平台为 Matlab 7.1.

仿真实验设定如下: (1) $\Phi \in R^{d \times L}$ 根据离散余弦变换生成, $\Phi = [\varphi_1, \dots, \varphi_L], \varphi_i \in R^{d \times 1}$, 原子个数 $L = 60, d = 60$; (2) 训练样本矩阵 $Y \in R^{d \times N}, N = 200; Y = [y_1, y_2, \dots, y_N], y_i = \sum_{l=1}^6 \text{coef}_l \varphi_{R_l} + \sigma_n \times \text{noise}$, noise 为零均值高斯随机噪声, 标准差 σ_n 分别取 [5, 10, 15]. 生成信号时满足: $\prod_{1 \leq l, j \leq 6} (R_l - R_j) \neq 0$, 原子的系数 $\text{coef}_l, 1 \leq l \leq 6$ 均通过随机数产生; (3) 待学习的字典稀疏矩阵为 $A \in R^{L \times K}$, 维数 $K = 80, D = \Phi A$.

取 $D^{[0]} = [y_1, \dots, y_K]$ 作为初始化字典, 使用修正的正交匹配追踪算法^[12]计算 A 的初始化 A_0 和 X 的初始化 X_0 . 然后分别应用稀疏性 K-SVD 和 SDLA-OSF 对样本进行稀疏性字典学习, 得到稀疏性字典. 其中, 稀疏性 K-SVD 算法的迭代次数设为 $K_T = 50$, 稀疏表示中, 稀疏度限制为 $T_X = 8, T_A = 8$. SDLA-OSF 参数设定为 $\lambda_X = \lambda_A = 5, K_X = K_A = 10, K_T = 50$.

4.1 算法收敛性对比

比较两种算法中目标函数值随迭代次数 t 的变化趋势. 为缩小待显示数值的范围, 在图中显示的是目标函数的对数值. 对稀疏性 K-SVD, 显示的为: $\log_{10}(\|Y - \Phi A X\|_F^2)$; 对于 SDLA-OSF, 显示的为: $\log_{10}(\|Y - \Phi A X\|_F^2 + \lambda_X J(X) + \lambda_A J(A))$.

无噪声是理想的, 实际中难以满足. 因此本文只考虑含噪情况. 对每种噪声水平重复十次实验, 取十次实验结果的目标函数的平均值作为该噪声情况下的目标函数. 以每种噪声水平下目标函数对数值的差分序列为观察对象. 将两种算法在三种噪声水平下的结果放在一起, 进行对比. 其结果如图 1 所示(为了能够显示清楚, 只选取了后面 20 步迭代的结果). 图中的三种噪声 noise1, noise2, noise3 分别对应标准差 5, 10, 15. “K-”表示 K-SVD 方法的结果, “S-”表示 SDLA-OSF 的结果.

从图 1 中可以看出, 对于 K-SVD 算法, 其三种噪声水平下的目标函数对数值的差分序列均与零轴有多次

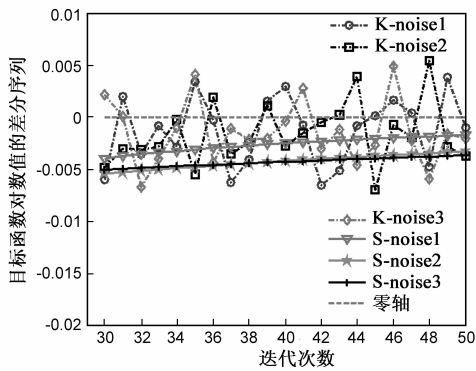


图1 噪声情况下目标函数对数值差分序列随迭代次数变化曲线交叉,这也意味着其目标函数在迭代过程中存在数次震荡,而三种噪声水平下 SDLA-OSF 算法对应的差分序列始终处于零轴下方,这表明其目标函数在迭代过程中是严格下降的.从而,在存在噪声情况下,K-SVD 算法迭代过程不是严格收敛的,而 SDLA-OSF 算法迭代过程是严格收敛的,这与理论分析的结论相符合.

4.2 收敛速度对比

给出的两种算法完成一次字典学习所耗费的时间如表 1 所示.(计算机配置:处理器:AMD Athlon (tm) 7550,主频:2.51GHz,内存:2.00GB).

表 1 两种算法运行时间对比(单位:秒)

标准差 \ 算法	稀疏性 K-SVD	SDLA-OSF
$\sigma_n = 0$	64.8594	7.625
$\sigma_n = 5$	61.6094	7.625
$\sigma_n = 10$	61.3594	7.6406
$\sigma_n = 15$	60.1094	7.5781

通过运行时间比较可以发现,稀疏性 K-SVD 耗费的时间约为 SDLA-OSF 耗费时间的 9 倍,也就是说 SDLA-OSF 的执行效率要优于稀疏性 K-SVD.

5 总结与讨论

针对稀疏性字典的学习问题,本文从整体优化角度出发,结合块松弛思想,经过严格数学推导,给出了 SDLA-OSF 算法,并进行了收敛性能分析.仿真实验结果表明 SDLA-OSF 在收敛性和执行效率方面均优于稀疏性 K-SVD 算法.

最后简单讨论一下 SDLA-OSF 目标函数中的两个参数 λ_X 、 λ_A .从算法 1 和算法 2 可知,这两个参数能影响收敛速度,如果取值较大,则会加速收敛,但使得字典学习的结果在对训练样本进行稀疏表示时误差较大,但若取值过大,则会使得 A 、 X 变成零矩阵(由式(10)和式(11)可以看出这点).在实际使用时,可先取较小的值,再做适当的调整.

参考文献

[1] 蔡泽民,赖剑煌.一种基于超完备字典学习的图像去噪方

法[J].电子学报,2009,37(2):347-350.

Cai Ze-min, Lai Jian-hui. An over-complete learned dictionary-based image de-noising method[J]. Acta Electronica Sinica, 2009, 37(2):347-350. (in Chinese)

- [2] M Aharon, M Elad, A M Bruckstein. The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation[J]. IEEE Transactions, 2006, 54(11):4311-4322.
- [3] M Yaghoobi, T Blumensath, M E Davies. Dictionary learning for sparse approximations with the majorization method[J]. IEEE Transactions, 2009, 57(6):2178-2191.
- [4] K Skretting, K Engan. Recursive least squares dictionary learning algorithm[J]. IEEE Transactions, 2010, 58(4):2121-2130.
- [5] B V Gowreesunker, A H Tewfik. Learning sparse representation using iterative subspace identification[J]. IEEE Transactions, 2010, 58(6):3055-3065.
- [6] N Dobigeon, J Y Tourneret. Bayesian orthogonal component analysis for sparse representation[J]. IEEE Transactions, 2010, 58(5):2675-2685.
- [7] R Rubinstein, M Zibulevsky, M Elad. Double sparsity: learning sparse dictionaries for sparse signal approximation[J]. IEEE Transactions, 2010, 58(3):1553-1564.
- [8] S Cotter, B Rao, K Engan, K K Delgado. Sparse solutions to linear inverse problems with multiple measurement vectors[J]. IEEE Transactions, 2005, 53(7):2477-2488.
- [9] 张贤达. 矩阵分析与应用[M]. 北京:清华大学出版社, 2004.270-271.
- [10] I Daubechies, M Defrise, C D Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint[J]. Comm Pure Appl Math, 2004, 57(11):1413-1441.
- [11] A Rantzer. On the dual of Lyapunov's second theorem[A]. Proceedings of the American Control Conference[C]. Chicago, America: AACC, 2000. 1186-1189.
- [12] D Needell. Topics in Compressed Sensing[D]. CA, Davis: University of California, 2009.

作者简介



高 磊 男,汉族,1983年2月生于安徽庐江,现为国防科技大学 ATR 重点实验室博士生,主要研究方向:稀疏信号处理、压缩感知及其在宽带雷达中的应用.

E-mail: ren_lgao@126.com

陈曾平 男,汉族,1967年10月生于福建福清,现为国防科技大学 ATR 重点实验室教授,博士,博士生导师.主要研究方向:信号获取、处理及其实现技术.E-mail: atrchen@sina.com