Signal Processing 114 (2015) 231–244

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Complex-valued sparse recovery via double-threshold sigmoid penalty

Shi Jun *,[1], Ding Renhuan [1], Xiang Gao [1], Zhang Xiaoling [1]

University of Electronic Science and Technology of China (UESTC), Chengdu, PR China

ARTICLE INFO

Article history:
Received 15 April 2014
Received in revised form
14 January 2015
Accepted 1 March 2015
Available online 10 March 2015

Keywords:
Convergence analysis
Double-threshold sigmoid penalty
Generalized iteratively reweighted least squares
Sparse recovery

ABSTRACT

The thresholding methods based on the generalized iteratively reweighted least squares (IRLS) iteration are discussed under the complex-valued condition in this paper. A new thresholding function (Double-Threshold Sigmoid (DTHS) function) and two associated algorithms (DTHS-1 and DTHS-2) are proposed herein, and their convergence performances are discussed in detail. It is shown that the generalized IRLS algorithm is unbiased if the thresholding penalty can eliminate the undesired perturbation term added on the correlation matrix of the measurement matrix. Compared with the others, the new algorithms are endowed with stability and insensitivity with respect to the regularization parameter by selecting some sound upper thresholds and dividing the iteration procedures into the degraded stage and DTHS stage respectively. Further analyses show that the DTHS-1 algorithm is suitable to deal with the sparse and continuous problems for both of the i.i.d. random matrix and under-resolution PSF matrix. The noise performance of the DTHS-1 algorithm is always superior to that of the IRLS algorithm, especially in the face of the under-resolution PSF matrix.

1. Introduction

In many fields, the signal recovery problem can be molded as solving a system of linear equations $\mathbf{Ax}=\mathbf{y}$, where, $\mathbf{A}$ is known as the measurement matrix or sensing matrix, $\mathbf{x}$ is the recovered vector and $\mathbf{y}$ is the observation vector. When the measurement matrix $\mathbf{A}$ is invertible, $\mathbf{x}$ can be solved via the inverse of $\mathbf{A}$; when the linear equations system is overdetermined, the least square (LS) solution is widely adopted. When the system of linear equations is underdetermined or ill-conditioned, however, there is more than one solution, and some further constraints are necessary.

The sparse recovery aims to find out the sparsest solution for an underdetermined problem, which is typically described as an optimization problem: $\min \|\mathbf{x}\|_0$, s.t., $\|\mathbf{Ax}-\mathbf{y}\|_2^2 \le \varepsilon$, where, $\|\mathbf{x}\|_0$ denotes the L0 (quasi) norm. Unfortunately, this optimization problem is proved to be NP-hard [1,2]. As a result, a series of endeavors to find out some suboptimal solutions are carried out in recent years.

The orthogonal matching pursuit (OMP) [3–8] approach heuristically solves L0 problem by iteratively reducing the residual error with a minimum increment of the nonzero components (supports) based on the greedy strategy, which can efficiently find the sparse solution. Further research shows that when the mutual coherence [9–11] or restricted isometry constant [12,13] is small and the recovered vector is sparse enough, the OMP algorithm can always obtain the desired sparse solution. Besides of the standard OMP, there are some modifications, such as the stagewise OMP [14], which have been researched in recent years.

Another popular strategy is approximating the L0 norm by L1 norm, i.e., $\min \|\mathbf{x}\|_0$, s.t., $\|\mathbf{Ax}-\mathbf{y}\|_2^2 \le \varepsilon$, which is known as the basis pursuit (BP) [11,15] algorithm. Since the L1 norm is convex, the optimization method was fully developed. By replacing the constraint $\|\mathbf{Ax}-\mathbf{y}\|_2^2 \le \varepsilon$ as $\|\mathbf{Ax}-\mathbf{y}\|_1 \le \varepsilon$ or $\|\mathbf{Ax}-\mathbf{y}\|_\infty \le \varepsilon$, the BP algorithm can be converted as a linear

* Corresponding author.
E-mail address: cn.papyrus@gmail.com (S. Jun).
[1] Tel.: +86 28 61830372.

http://dx.doi.org/10.1016/j.sigpro.2015.03.001
0165-1684/

programming problem and solved by many optimization tools. Further research shows that when the problem is sparse enough, the unique solution of the BP algorithm is identical to the solution of L0 problem [7].

The third strategy to find the sparsest solution is the regularization technique that converts the optimization problem to an unconstrained optimization by treating the L1 norm as a penalty term. In this case, the sparsest solution can be obtained by some gradient-based methods, such as the iteratively reweighted least squares (IRLS) [7,16,17]. Besides of the L1 norm, some other penalties (named as the thresholding methods), including the FOCUSS method [18] using $\|\mathbf{x}\|_\gamma$ as the penalty, the Mangasarian's penalty [19], and the MC+ penalty [20,21], are applied to approximate the L0 norm more accurately. Since those modified penalties are non-convex, the iteration sequences might trap into some undesired local minima.

In this paper, the thresholding methods based on the generalized IRLS iteration are discussed, and a novel penalty function, double-threshold sigmoid (DTHS) penalty, is proposed and analyzed in detail. The complex-valued problem is also considered due to that in many fields, such as radar/sonar detection [22–24], synthetic aperture radar (SAR) image enhancement [25–31], the signal models are always complex-valued, and it is tedious to convert them into some real-valued counterparts and solve them by using a real-valued-oriented method. Since the complex-valued problem is completely compatible to the real-valued one, the conclusions drawn in this case can also be used for the real-valued problem.

The organization of this paper is as follows. In Section 2, the iteratively reweighted least squares algorithm is introduced briefly. In Section 3, some typical thresholding penalties are reviewed, a new thresholding penalty, DTHS penalty, is proposed, and its convergence performance is discussed in detail. In Section 4, the performances of the DTHS penalty are analyzed by some numerical experiments and the effects of the regularization parameter on the different methods are compared. In Section 5, the performances of the OMP, BP, IRLS and DTHS-1 methods are compared in both of the i.i.d. random matrix and under-resolution PSF matrix cases. A summary is presented in Section 6 finally.

2. Review on L1 regularization

The L1 regularization method attempts to find the sparsest solution based on the penalty strategy, whose model can be written as

$$\min_{\mathbf{x}} J(\mathbf{x};\lambda) \tag{1.1}$$

$$J(\mathbf{x};\lambda)=\frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2+\lambda\|\mathbf{x}\|_1 \tag{1.2}$$

where $\|\mathbf{x}\|_1$ is treated as the penalty term, λ denotes the regularization parameter.

As we all know, the optimal solution satisfies the first-order optimization conditions:

$$\nabla_{\mathbf{x}} J(\mathbf{x}^*;\lambda)=0 \tag{2}$$

An obstacle before deducing the gradient of Eq. (1.2) is that the derivative of $|x|$ is undefined at zero, which always happens when dealing with the sparse problem. Thus, some approximation tricks, named as the smoothing technique [32], are adopted in practice. In the real-valued case, two frequently-used approximations are

$$|x|\approx\sqrt{x^2+\delta} \tag{3.1}$$

$$|x|\approx\frac{1}{\alpha}\left[\ln(1+e^{-\alpha x})+\ln(1+e^{\alpha x})\right] \tag{3.2}$$

where δ is a small positive constant, Eq. (3.1) approaches $|x|$ when $\delta\to 0^+$; α is a positive constant, Eq. (3.2) approaches $|x|$ when $\alpha\to\infty$. The former approximation is used in this paper and δ is fixed as 0.0001.

With Eq. (3.1), the derivative of $|x|$ can be approximated as

$$\frac{d|x|}{dx}\approx\frac{x}{\sqrt{x^2+\delta}} \tag{4}$$

In the complex-valued case, we need to calculate the derivative with respect to the real part x_r and the imaginary part x_i respectively:

$$\frac{\partial\sqrt{x_r^2+x_i^2+\delta}}{\partial x_r}\approx\frac{x_r}{\sqrt{x_r^2+x_i^2+\delta}} \tag{5.1}$$

$$\frac{\partial\sqrt{x_r^2+x_i^2+\delta}}{\partial x_i}\approx\frac{x_i}{\sqrt{x_r^2+x_i^2+\delta}} \tag{5.2}$$

Obviously, ensuring the real and imaginary parts to be equal to zero respectively is equivalent to making the following formula equal to zero, i.e.

$$\frac{x_r}{\sqrt{x_r^2+x_i^2+\delta}}+\frac{\sqrt{-1}x_i}{\sqrt{x_r^2+x_i^2+\delta}}=\frac{x}{\sqrt{|x|^2+\delta}}=0 \tag{6}$$

where x is complex-valued. Eq. (6) indicates that Eq. (4) still holds in the complex-valued case.

With the above preparations, the gradient of Eq. (1.2) can be written as

$$\nabla_{\mathbf{x}} J(\mathbf{x};\lambda)=\widetilde{\mathbf{H}}_1\mathbf{x}-\mathbf{A}^H\mathbf{y} \tag{7.1}$$

$$\widetilde{\mathbf{H}}_1=\mathbf{A}^H\mathbf{A}+\lambda\mathbf{X}_1 \tag{7.2}$$

$$\mathbf{X}_1\triangleq diag\left(1/\sqrt{|x_l|^2+\delta}\right) \tag{7.3}$$

where $\mathbf{X}_1$ is a diagonal matrix, whose diagonal entries are $1/\sqrt{|x_l|^2+\delta}$.

To accelerate the iteration process, we approximate the Hessian matrix of Eq. (1.2) as

$$\nabla_{\mathbf{x}}^2 J(\mathbf{x};\lambda)\approx\widetilde{\mathbf{H}}_1 \tag{8}$$

and obtain the quasi Newton iteration formula of the L1 regularization problem, which is written as

$$\begin{aligned}\mathbf{x}^{[k+1]}&=\mathbf{x}^{[k]}-h_{\mathbf{x}}\widetilde{\mathbf{H}}_1^{-1}\left(\widetilde{\mathbf{H}}_1\mathbf{x}^{[k]}-\mathbf{A}^H\mathbf{y}\right)\\&=(1-h_{\mathbf{x}})\mathbf{x}^{[k]}+h_{\mathbf{x}}\widetilde{\mathbf{H}}_1^{-1}\mathbf{A}^H\mathbf{y}\end{aligned} \tag{9.1}$$

Eq. (9.1) is known as the iteratively reweighted least squares (IRLS) algorithm [7]. Just as it is suggested in literature [27], by fixing the step size $h_{\mathbf{x}}$ as 1, the iteration