SCIENCE CHINA
Information Sciences

• LETTER •

March 2022, Vol. 65 139303:1–139303:2
https://doi.org/10.1007/s11432-020-3126-7

Sparse signal reconstruction via generalized two-stage thresholding

Heping SONG[1,2*], Zehong AI[1], Yuping LAI[3*], Hongying MENG[4] & Qirong MAO[1]

[1]School of Computer Science and Telecommunications Engineering, Jiangsu University, Zhenjiang 212012, China;
[2]Jiangsu Key Laboratory of Security Technology for Industrial Cyberspace, Jiangsu University, Zhenjiang 212012, China;
[3]School of Cyberspace Security, Beijing University of Posts and Telecommunications, Beijing 100875, China;
[4]Department of Electronic & Computer Engineering, Brunel University London, Middlesex UB8 3PH, UK

Received 28 September 2020/Revised 12 November 2020/Accepted 24 November 2020/Published online 14 February 2022

Citation Song H P, Ai Z H, Lai Y P, et al. Sparse signal reconstruction via generalized two-stage thresholding. Sci China Inf Sci, 2022, 65(3): 139303, https://doi.org/10.1007/s11432-020-3126-7

Dear editor,

The sparse representation model has received a great amount of attention in various signal and image processing applications. Compressed sensing (CS) [1] has consistently focused on devising sparse representation methods that seek to efficiently reconstruct a k-sparse (only k nonzero entries, $k \ll n$) underlying n-length signal via a much smaller number of compressed measurements of length-m ($m \ll n$). Given an unknown k-sparse signal $x \in \mathbb{R}^n$, the sparse reconstruction problem is closely linked to solving a highly undetermined linear regression model with sparsity-constrained optimization,

$$\min_x \|x\|_0 \quad \text{s.t.} \quad y = Ax, \tag{1}$$

where x is the unknown signal to be reconstructed, $\|x\|_0$ is a measure of sparsity involving the number of nonzero entries in x, $y \in \mathbb{R}^m$ is the measurement vector, and $A \in \mathbb{R}^{m\times n}$ is the sampling/measurement matrix. The core insight in the seminal paper on CS [1] is to solve ℓ_1 minimization problems for (1) instead of ℓ_0 minimization problems owing to NP-hardness (nonconvexity and combinatorial effects). This is the basis pursuit (BP) problem in [2]:

$$\min_x \|x\|_1 \quad \text{s.t.} \quad y = Ax. \tag{2}$$

Subsequently, researchers have presented various numerical approaches for sparse signal reconstruction [3]. In general, there are two major classes of numerical approaches: convex relaxation and greedy algorithms. Convex relaxation uses ℓ_1 norm of x in (2) instead of the ℓ_0 norm in (1). In 2010, an algorithmic framework called iterative support detection (ISD) was presented in [4]; this framework improves the failed BP reconstructions by detecting the support set through the solution of a truncated ℓ_1 optimization problem (2). Another line of research concerns greedy algorithms such as iterative hard thresholding, hard thresholding gradient descent (GraDes) [5], or hard thresholding pursuit (HTP) [6] for approximating the ℓ_0-constrained solution to (1). The key part of greedy algorithms is to iteratively improve a sparse solution by successively selecting one or more elements at a time. Yuan et al. [7] presented gradient hard thresholding pursuit (GraHTP) as a generalization of HTP from CS to the generic problem of sparsity-exploiting loss functions. Huang et al. [8] proposed a numerical method called support detection and root finding (SDAR) to approximate (1) motivated from the KKT conditions for the ℓ_0-constrained least squares problem. The key idea is to iteratively create a sequence of solutions based on support detection by referencing the primal and dual information and root finding using truncated least squares optimization.

In this study, inspired by the success of ISD and SDAR in sparse signal reconstruction, we propose a novel algorithmic framework, called generalized two-stage thresholding (GTST), to determine sparse solutions of undetermined linear systems. GTST regards support detection as the first stage thresholding, which detects a support set I using an approximate iteration of (1) as the reference, and signal estimation as the second stage thresholding, which estimates a new reconstruction by solving a truncated ℓ_0-constrained minimization problem on the detected support I; it repeats these two stages for a few iterations. GTST generalizes ISD from BP based convex relaxation to a generic ℓ_0-constrained minimization problem setup of sparsity-exploiting convex optimization. It means that GTST exhibits more flexible thresholding methods than SDAR. In addition, we introduce an efficient implementation of GTST (dubbed as GTST-α) equipped with an effective thresholding method for support detection. GTST-α runs as fast as SDAR but requires significantly fewer measurements. GTST-α is sparsity adaptive by introducing an easy-to-tune parameter α. This is an appealing feature when the sparsity level is unknown. In contrast to ISD, GTST-α is easy-to-implement.

Methods. Considering a CS sparse signal reconstruction problem, the greedy algorithms aim to detect the informative support I and estimate y by k-term approximation

* Corresponding author (email: songhp@ujs.edu.cn, laiyuping@qq.com)

through $y = A_I x_I$. Then, reconstruct the unknown sparse signal x by solving the truncated version of the least squares problem,

$$x_I = \arg\min_z \{\|y - A_I z\|_2^2,\ \mathrm{supp}(z) \subseteq I\}, \quad x_T = \mathbf{0}, \quad (3)$$

where the matrix A_I refers to the restricting columns from A indexed by I, $T = [1,\ n] \setminus I$ is the complement set of I on $\{1, 2, \ldots, n\}$. Let $\mathrm{supp}(x)$ denote the support set of x; i.e., the index set corresponds to the nonzero elements in x. The approximate analytical solution of x_I is $A_I^\dagger y = (A_I' A_I)^{-1} A_I' y$. For large-scale applications, the truncated least squares problem can be easily implemented using the conjugate gradient method. We define a general class of greedy algorithms, termed GTST. Starting from a solution $x = \mathbf{0}$, GTST iterates between two main steps.

Step 1: Support detection. Identify the support set I of the unknown signal x by referencing the step of approximating solution of (1).

Step 2: Signal reconstruction. Reconstruct the true sparse signal x by solving the truncated version of least squares problem, $x_I = A_I^\dagger y$, $x_T = \mathbf{0}$.

Like ISD, GTST is an algorithmic framework. For GTST, successful sparse signal reconstruction depends on support detection, which requires an efficient reference and an effective support detection strategy. We employ the GraDes step [5] as the reference for support detection and introduce an efficient implementation of GTST equipped with a dynamic thresholding strategy. The complete description of the proposed algorithms is presented here.

Step 1: Initialization. Initialize primal approximation $x^0 = \mathbf{0}$, residual signal $r^0 = y$, dual approximation $d^0 = A' r^0$, support set $I^0 = \emptyset$, complement set $T^0 = [1, n]$, and set the iteration counter $t = 1$.

Step 2: Support detection. Update signal approximation: $x^t = x^{t-1} + d^{t-1}$.

Detect the support set I: $I^t = \{i : |x_i^t| \geqslant \alpha\ \max |x_j^t|;\ i \in [1, n],\ j \in T^{t-1},\ \alpha \in (0, 1]\}$, $T^t = [1, n] \setminus I^t$.

Step 3: Signal reconstruction. Estimate the signal: $x_{I^t}^t = A_{I^t}^\dagger y$, $x_{T^t}^t = \mathbf{0}$.

Update the residual: $r^t = y - A x^t$.

Update the dual approximation: $d_{T^t}^t = A_{T^t}' r^t$, $d_{I^t}^t = \mathbf{0}$.

Step 4: Halting. Check whether the stopping condition is false. If so, update $t = t + 1$ and back to Step 2. Otherwise, the reconstructed signal x has nonzero elements in the support set I^t and $x_{I^t}^t$ corresponds to the support vector.

We have presented the implementation of GTST, which employs the GraDes step as the reference for support detection, and the corresponding algorithm is referred to as GTST-α (as shown in Appendix A), where $\alpha \in (0, 1]$ is the thresholding parameter. The key idea is to constrain the GraDes step to be sparse via dynamic thresholding. GTST-α iteratively refines the detected support I, which is not necessarily nested or increasing over the iterations. However, SDAR needs a hard thresholding to fix the cardinality of the support set I and ensures that the new approximation is k-sparse. The most appealing feature of GTST-α is its capability of sparse recovery without prior knowledge about the underlying sparsity level k. This makes it a promising candidate for many practical applications. GraDes can be considered as a single-stage thresholding method, which does not involve a refining step on the detected support set I. GTST updates the reconstructed signal using a truncated least-squares solution on the detected support set I with size no more than k. However, ISD progressively refines the ℓ_1 solution on the complement set of the support set I with larger sizes.

Experiments. We evaluate the sparse reconstruction ability of GTST and compare it with GraDes and SDAR, where the latter is a state-of-the-art algorithm. First, we present a demo of signal reconstruction performed in [4]. GTST-α can perfectly reconstruct a sparse signal with 25 nonzeros using only 60 measurements with signal dimension $n = 200$, but SDAR fails to do so under the same setting. Second, we demonstrate the sparse reconstruction performance of GTST in terms of the probability of exact reconstruction. GTST-α obtains better reconstruction performance than GraDes and SDAR. For more detailed experimental results, please refer to Appendixes A–C.

Conclusion. In this work, we proposed a novel algorithmic framework, dubbed as GTST, for CS sparse signal reconstruction. GTST alternates between detecting a support set of the unknown true signal by referencing the step of approximating solution of ℓ_0-constrained least-squares problem and reconstructing the sparse signal by solving a truncated version of least-squares optimization on the detected support. Moreover, we proposed an efficient implementation of GTST equipped with an effective thresholding approach for support detection. The experimental studies demonstrated that GTST outperforms SDAR and GraDes in terms of the probability of exact reconstruction.

Acknowledgements This work was supported by Beijing Municipal Natural Science Foundation (Grant No. 4194076), National Natural Science Foundation of China (Grant Nos. U1836220, 61672267), Jiangsu Province Natural Science Foundation (Grant No. BK20170558), and China Scholarship Council (Grant No. 202008320094).

Supporting information Appendixes A–C. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Donoho D L. Compressed sensing. IEEE Trans Inform Theor, 2006, 52: 1289–1306
2 Chen S S, Donoho D L, Saunders M A. Atomic decomposition by basis pursuit. SIAM Rev, 2001, 43: 129–159
3 Wang Y, Meng D, Yuan M. Sparse recovery: from vectors to tensors. Natl Sci Rev, 2018, 5: 756–767
4 Wang Y, Yin W. Sparse signal reconstruction via iterative support detection. SIAM J Imag Sci, 2010, 3: 462–491
5 Garg R, Khandekar R. Gradient descent with sparsification: an iterative algorithm for sparse recovery with restricted isometry property. In: Proceedings of the 26th Annual International Conference on Machine Learning, 2009. 337–344
6 Foucart S. Hard thresholding pursuit: an algorithm for compressive sensing. SIAM J Numer Anal, 2011, 49: 2543–2563
7 Yuan X, Li P, Zhang T. Gradient hard thresholding pursuit. J Mach Learn Res, 2018, 18: 1–43
8 Huang J, Jiao Y, Liu Y, et al. A constructive approach to L0 penalized regression. J Mach Learn Res, 2018, 19: 1–37