Signal Processing 144 (2018) 265–270

Contents lists available at ScienceDirect

Signal Processing

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

Short communication

AKRON: An algorithm for approximating sparse kernel reconstruction

Gregory Ditzler[a,*], Nidhal Carla Bouaynaya[b], Roman Shterenberg[c]

[a] Department of Electrical and Computer Engineering, The University of Arizona, Tucson, AZ 85721, USA
[b] Department of Electrical and Computer Engineering, Rowan University, Glassboro, NJ 08028, USA
[c] Department of Mathematics, University of Alabama at Birmingham, Birmingham, AL 35294, USA

ARTICLE INFO

Article history:
Received 17 April 2017
Revised 14 September 2017
Accepted 13 October 2017
Available online 17 October 2017

Keywords:
Compressive sensing
Linear systems
Optimization

ABSTRACT

Exact reconstruction of a sparse signal for an under-determined linear system using the ℓ_0-measure is, in general, an NP-hard problem. The most popular approach is to relax the ℓ_0-optimization problem to an ℓ_1-approximation. However, the strength of this convex approximation relies upon rigid properties on the system, which are not verifiable in practice. Greedy algorithms have been proposed in the past to speed up the optimization of the ℓ_1 problem, but their computational efficiency comes at the expense of a larger error. In an effort to control error and complexity, this paper goes beyond the ℓ_1-approximation by growing neighborhoods of the ℓ_1-solution that moves towards the optimal solution. The size of the neighborhood is tunable depending on the computational resources. The proposed algorithm, termed Approximate Kernel RecONstruction (AKRON), yields significantly smaller errors than current greedy methods with a controllable computational cost. By construction, the error of AKRON is smaller than or to equal the ℓ_1-solution. AKRON enjoys all the error bounds of ℓ_1 under the restricted isometry property condition. We benchmarked AKRON on simulated data from several under-determined systems, and the results show that AKRON can significantly improve the reconstruction error with slightly more computational cost than solving the ℓ_1 problem directly.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Many engineering problems are formulated as inverse problems, which is where the number of parameters (p) greatly exceeds the number of measurements (n) available. Examples include: source estimation of electroencephalographic (EEG) and magnetoencephalographic (MEG) data [1,2], reverse-engineering of genetic regulatory networks from high-throughput gene expression data [3,4], magnetic resonance imaging [5], information theory and communication engineering [6], and electromagnetics and antenna design [7]. These inverse problems, known as "large p small n", pose a challenge, because of the non-identifiability of a solution. Additional constraints or prior knowledge are needed to solve such under-determined systems. In many cases, such as inference of genetic regulatory networks [3,4], we are interested in the sparsest solution. The objective is then to recover the sparsest signal, $\mathbf{x} \in \mathbb{C}^p$, from a measurement matrix, $\mathbf{A} \in \mathbb{C}^{n \times p}$, and observed vector $\mathbf{y} \in \mathbb{C}^n$ such that $\mathbf{y} = \mathbf{A}\mathbf{x}$, where $n \ll p$. In a noisy setting, the problem is formulated as $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$, where $\mathbf{e}$ is a vector of measurement noise with a bounded variance, i.e., $\|\mathbf{e}\|_2 \leq \epsilon$. Without loss of generality, it is assumed that $\mathbf{A}$ is full-rank; otherwise, the observations would be redundant.

Finding the sparsest solution amounts to solving the following optimization problem:

$$\mathbf{x}^* = \text{argmin}\{\|\mathbf{x}\|_0 : \mathbf{A}\mathbf{x} = \mathbf{y}\}, \tag{1}$$

where $\|\mathbf{x}\|_0$ denotes the ℓ_0-measure of vector $\mathbf{x}$, i.e., the number of non-zero elements of $\mathbf{x}$. Observe that ℓ_0 is not a proper norm and that is why we refer to it as a "measure" although, by abuse of notation, we may also write ℓ_0-norm. Unfortunately, (1) is in general an NP-hard combinatorial problem since it involves finding the number *and* positions of the zeros in a p-dimensional space [8]. The field of compressive sensing (CS) addresses this problem by solving the under-determined system with a unique sparsest solution under specific conditions on the system. The ℓ_0-norm objective in (1) can be relaxed to the ℓ_1-norm, solving the following convex optimization problem:

$$\widehat{\mathbf{x}}_1 = \text{argmin}\{\|\mathbf{x}\|_1 : \mathbf{A}\mathbf{x} = \mathbf{y}\}. \tag{2}$$

This convex relaxation makes the problem more tractable; however, in general, the solutions of (1) and (2) are not equivalent. CS theory has shown that, if $\mathbf{A}$ satisfies the null space property (NSP) or the restricted isometry property (RIP), then the ℓ_1 problem yields the optimal ℓ_0 solution [8]. Unfortunately, these conditions

* Corresponding author.
E-mail addresses: ditzler@email.arizona.edu (G. Ditzler), bouaynaya@rowan.edu (N. Carla Bouaynaya), shterenb@uab.edu (R. Shterenberg).

https://doi.org/10.1016/j.sigpro.2017.10.020
0165-1684/© 2017 Elsevier B.V. All rights reserved.

are not verifiable in practice. In particular, one cannot check if the obtained ℓ_1-solution is the sparsest solution or not! Through examples and simulations, we show that, in general, the ℓ_1-solution may be far from the ℓ_0-optimal solution. Hence, it is crucial to develop greedy algorithms that achieve a balance between computational complexity and reconstruction error.

2. Related work

Recent efforts focused on greedy algorithms to infer a sparse solution. In particular, a family of *Hard Thresholding* (HT) algorithms have been suggested in [9], which makes an initial guess for the support of $\mathbf{x}$ and then projects the measurements $\mathbf{y}$ onto this support. An iterative version called *Iterative Hard Thresholding* (IHT) updates the residual and estimates a new $\mathbf{x}$ at every iteration until the stopping criterion is satisfied. Another version of lower complexity per iteration, called *Matching Pursuit* (MP), has been suggested. The *Orthogonal Matching Pursuit* (OMP) [10] is an iterative greedy algorithm that selects at each step the column which is most correlated with the current residuals, and estimates the nonzero entries in the vector $\mathbf{x}$ with a computational complexity $O(k \log p)$. OMP's computational improvements, however, come at the cost of increased reconstruction error. *Compressive Sampling Matching Pursuit* (CoSaMP) [10] combines the approaches of OMP and HT in a two-stage greedy algorithm that aims to improve the reconstruction error of OMP. Unfortunately, these methods sacrifice accuracy of the reconstruction for the runtime as they approximate the ℓ_0-norm by other cost functions. Recently, SL0, or *smoothed* ℓ_0, has been proposed as a fast algorithm to directly approximate the ℓ_0 solution [11]. Candés et al. [5] proposed an iterative re-weighted ℓ_1 minimization algorithm that has theoretical guarantees that is can improve the ℓ_1 solution [12].

In our previous work, we presented *Kernel RecONstruction* (KRON), a greedy algorithm, that achieves an exact solution to (1), without exhaustively searching $\mathbb{C}^p$ [13]. In KRON, finding the sparsest solution amounts to solving $\binom{p}{s=p-n}$ linear equations. All $\binom{p}{s}$ potential solutions have at least s zeros. The sparsest solution is guaranteed to be one of them. The computational complexity of KRON is $\mathcal{O}(p^s)$. KRON yields the optimal sparsest solution (zero reconstruction error) at a high computational cost.

Against this background, we seek to develop an approach for approximating (1) yielding reconstruction errors lower than ℓ_1-norm, and other approaches such as OMP and CoSaMP, and at the same time having comparable, or at least controllable, computational cost.

3. Approximate Kernel Reconstruction

3.1. Central idea behind AKRON

In this section, we motivate the central idea behind AKRON, given general linear algebra knowledge about the underdetermined system. First, we know that the system $\mathbf{Ax} = \mathbf{y}$ always admits solutions with $s = (p - n)$ zeros because s is the dimension of the Kernel of $\mathbf{A}$; hence the name Kernel RecONstruction (KRON) in [13]. KRON distributes s zeros among the p entries then searches for all the solutions with exactly s zeros. The sparsest solution is guaranteed to be among these $\binom{p}{s}$ solutions. However, we do not know in advance which one it will be. KRON tries all possible $\binom{p}{s}$ solutions then chooses the sparsest. Notice that no conditions are imposed on the matrix $\mathbf{A}$; that is, KRON recovers the optimal sparsest solution whether the RIP condition is satisfied or not. The central issue with KRON is that it becomes computationally prohibitive when p is large *and* s in the order of $\frac{p}{2}$. Therefore, we propose AKRON to reduce the number of enumerations that need to be performed in KRON. To achieve this, the central idea behind AKRON is to use the standard ℓ_1-approximation to "guess" the locations of the s zeros that will result in the sparsest solution. Finding s correct zero locations is *sufficient* to find the optimal sparsest solution. This idea can also be viewed as *a "perturbation" of the ℓ_1-approximation to make it closer to the ℓ_0-norm*. Formally, we define a δ-neighborhood of the ℓ_1-approximation that allows AKRON to find sparser solutions and reduce the reconstruction error. The size of the neighborhood is tunable depending on the computational power available, and vary from 0 (ℓ_1-approximation) to n (KRON, i.e., perfect reconstruction). In particular, when the ℓ_1-approximation is optimal (RIP conditions satisfied), AKRON is also optimal, but when the ℓ_1-solution is suboptimal, AKRON results in a better (i.e., sparser) solution with smaller recovery error.

3.2. The noiseless case

AKRON begins by solving the ℓ_1 convex optimization problem in (2). Denote the solution by $\widehat{\mathbf{x}}_1$. In general, $\widehat{\mathbf{x}}_1$ is different from the desired ℓ_0-solution. However, since ℓ_1 is the closest convex norm to ℓ_0, we can use $\widehat{\mathbf{x}}_1$ to find the locations of s zeros, which would correspond to the s-smallest magnitudes in $\widehat{\mathbf{x}}_1$. The central idea behind AKRON's 0-neighborhood solution is as follows: (1) find the indices ($\mathcal{Q}$) with the s-smallest magnitudes of the ℓ_1 solution, (2) set these indices to zero, then (3) re-solve the system $\mathbf{Ax} = \mathbf{y}$. Call this solution $\widehat{\mathbf{x}}^*_{\delta=0}$. The following proposition bounds the error between the ℓ_1-solution and the $(\delta = 0)$-neighborhood solution $\widehat{\mathbf{x}}^*_{\delta=0}$.

Proposition 1. *Let $\widehat{\mathbf{x}}_1$ denote the ℓ_1-solution of the under-determined problem in (2). Without loss of generality, we assume that $\mathbf{A} \in \mathbb{C}^{n\times p}$ is full-rank, and call $s = p - n$. Let $\{|\widehat{\mathbf{x}}_1(i_1)|, \ldots, |\widehat{\mathbf{x}}_1(i_s)|\}$ be the set of the s-smallest magnitudes of $\widehat{\mathbf{x}}_1$. Then, we have*

$$\|\widehat{\mathbf{x}}_1 - \widehat{\mathbf{x}}^*_{\delta=0}\|_2 \leq \sqrt{s}\, C_A \max\{|\widehat{\mathbf{x}}_1(i_1)|, \cdots, |\widehat{\mathbf{x}}_1(i_s)|\}, \tag{3}$$

where C_A is a constant that depends only on the matrix $\mathbf{A}$: $C_A = (1 + \|\mathbf{A}_Q^{-1}\|_2 \|\mathbf{A}_{Q^\perp}\|_2)$, where $\mathbf{A}_Q$ is the $(n \times n)$ sub-matrix of $\mathbf{A}$ obtained by removing the s columns indexed by $\{i_1, \cdots, i_s\}$, and $\mathbf{A}_{Q^\perp} \in \mathbb{C}^{n\times s}$ is the complement matrix, i.e., the matrix that contains only the columns corresponding to these s-smallest elements.

Proof. Denote by $\mathbf{A}_Q \in \mathbb{C}^{n\times n}$ the reduced matrix, where the columns corresponding to the indices of the s-smallest elements in $\widehat{\mathbf{x}}_1$ were removed. Notice that $\mathbf{A}_Q$ is invertible because $\mathbf{A}$ is full-rank. Let $\mathbf{A}_{Q^\perp} \in \mathbb{C}^{n\times s}$ be the complement matrix, i.e., the matrix that contains only the columns corresponding to the s-smallest elements $\{i_1, \cdots, i_s\}$. We adopt similar notations for $\widehat{\mathbf{x}}_{1_Q} \in \mathbb{C}^{n\times 1}$ and $\widehat{\mathbf{x}}_{1_{Q^\perp}} \in \mathbb{C}^{s\times 1}$. We have

$$\mathbf{A}\widehat{\mathbf{x}}_1 = \mathbf{A}\widehat{\mathbf{x}}^*_{\delta=0} = \mathbf{y} \tag{4}$$

$$\Longleftrightarrow$$

$$\mathbf{A}_Q\widehat{\mathbf{x}}_{1_Q} + \mathbf{A}_{Q^\perp}\widehat{\mathbf{x}}_{1_{Q^\perp}} = \mathbf{A}\widehat{\mathbf{x}}^*_{\delta=0}. \tag{5}$$

Observe that since, by construction, $\widehat{\mathbf{x}}^*_{\delta=0_{Q^\perp}} = \mathbf{0}$, we have that

$$\mathbf{A}_Q\widehat{\mathbf{x}}^*_{\delta=0_Q} = \mathbf{A}\widehat{\mathbf{x}}^*_{\delta=0}. \tag{6}$$

From Eqs. (5) and (6), we have

$$\mathbf{A}_Q(\widehat{\mathbf{x}}^*_{\delta=0_Q} - \widehat{\mathbf{x}}_{1_Q}) = \mathbf{A}_{Q^\perp}\,\widehat{\mathbf{x}}_{1_{Q^\perp}}.$$

Therefore,

$$\widehat{\mathbf{x}}^*_{\delta=0_Q} - \widehat{\mathbf{x}}_{1_Q} = \mathbf{A}_Q^{-1}\mathbf{A}_{Q^\perp}\,\widehat{\mathbf{x}}_{1_{Q^\perp}}.$$

Using norm inequalities, we obtain

$$\|\widehat{\mathbf{x}}^*_{\delta=0_Q} - \widehat{\mathbf{x}}_{1_Q}\|_2 \leq \|\mathbf{A}_Q^{-1}\|_2 \|\mathbf{A}_{Q^\perp}\|_2\, \|\widehat{\mathbf{x}}_{1_{Q^\perp}}\|_2. \tag{7}$$

Download English Version:

https://daneshyari.com/en/article/6958023

Download Persian Version:

https://daneshyari.com/article/6958023

Daneshyari.com