Digital Signal Processing 82 (2018) 91–105

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Accelerated orthogonal least-squares for large-scale sparse reconstruction

Abolfazl Hashemi *, Haris Vikalo

Department of Electrical and Computer Engineering, University of Texas at Austin, USA

A R T I C L E I N F O

Article history:
Available online 1 August 2018

Keywords:
Sparse recovery
Compressed sensing
Orthogonal least-squares
Orthogonal matching pursuit
Subspace clustering

A B S T R A C T

We study the problem of inferring a sparse vector from random linear combinations of its components. We propose the Accelerated Orthogonal Least-Squares (AOLS) algorithm that improves performance of the well-known Orthogonal Least-Squares (OLS) algorithm while requiring significantly lower computational costs. While OLS greedily selects columns of the coefficient matrix that correspond to non-zero components of the sparse vector, AOLS employs a novel computationally efficient procedure that speeds up the search by anticipating future selections via choosing L columns in each step, where L is an adjustable hyper-parameter. We analyze the performance of AOLS and establish lower bounds on the probability of exact recovery for both noiseless and noisy random linear measurements. In the noiseless scenario, it is shown that when the coefficients are samples from a Gaussian distribution, AOLS with high probability recovers a k-sparse m-dimensional sparse vector using $\mathcal{O}(k \log \frac{m}{k+L-1})$ measurements. Similar result is established for the bounded-noise scenario where an additional condition on the smallest nonzero element of the unknown vector is required. The asymptotic sampling complexity of AOLS is lower than the asymptotic sampling complexity of the existing sparse reconstruction algorithms. In simulations, AOLS is compared to state-of-the-art sparse recovery techniques and shown to provide better performance in terms of accuracy, running time, or both. Finally, we consider an application of AOLS to clustering high-dimensional data lying on the union of low-dimensional subspaces and demonstrate its superiority over existing methods.

© 2018 Elsevier Inc. All rights reserved.

1. Introduction

The task of estimating sparse signal from a few linear combinations of its components is readily cast as the problem of finding a sparse solution to an underdetermined system of linear equations. Sparse recovery is encountered in many practical scenarios, including compressed sensing [1], subspace clustering [2,3], sparse channel estimation [4,5], compressive DNA microarrays [6], and a number of other applications in signal processing and machine learning [7–9]. Consider the linear measurement model

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \boldsymbol{\nu}, \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^n$ denotes the vector of observations, $\mathbf{A} \in \mathbb{R}^{n\times m}$ is the coefficient matrix (i.e., a collection of features) assumed to be full rank (generally, $n < m$), $\boldsymbol{\nu} \in \mathbb{R}^n$ is the additive measurement noise vector, and $\mathbf{x} \in \mathbb{R}^m$ is an unknown vector assumed to have at most k non-zero components (i.e., k is the sparsity level of $\mathbf{x}$). Finding a sparse approximation to $\mathbf{x}$ leads to a cardinality-constrained least-squares problem

$$\underset{\mathbf{x}}{\text{minimize}} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{subject to} \quad \|\mathbf{x}\|_0 \le k, \tag{2}$$

known to be NP-hard; here $\|\cdot\|_0$ denotes the ℓ_0-norm, i.e., the number of non-zero components of its argument. The high cost of finding the exact solution to (2) motivated development of a number of heuristics that can generally be grouped in the following categories:

1) *Convex relaxation schemes.* These methods perform computationally efficient search for a sparse solution by replacing the non-convex ℓ_0-constrained optimization by a sparsity-promoting ℓ_1-norm optimization. It was shown in [10] that such a formulation enables exact recovery of a sufficiently sparse signal from noise-free measurements under certain conditions on $\mathbf{A}$ and with $\mathcal{O}(k \log \frac{m}{k})$ measurements. However, while the convexity of ℓ_1-norm enables algorithmically straightforward sparse vector recovery by means of, e.g., iterative shrinkage-thresholding [11] or alternating direction method of multipliers [12], the complexity of such methods is often prohibitive in settings where one deals with high-dimensional signals.

* Corresponding author.
E-mail address: abolfazl@utexas.edu (A. Hashemi).

https://doi.org/10.1016/j.dsp.2018.07.010
1051-2004/© 2018 Elsevier Inc. All rights reserved.

2) *Greedy schemes.* These heuristics attempt to satisfy the cardinality constraint directly by successively identifying k columns of the coefficient matrix which correspond to non-zero components of the unknown vector. Among the greedy methods for sparse vector reconstruction, the orthogonal matching pursuit (OMP) [13] and Orthogonal Least-Squares (OLS) [14,15] have attracted particular attention in recent years. Intuitively appealing due to its simple geometric interpretation, OMP is characterized by high speed and competitive performance. In each iteration, OMP selects a column of the coefficient matrix $\mathbf{A}$ having the highest correlation with the so-called residual vector and adds it to the set of active columns; then by solving a least-square problem using the modified Gram–Schmidt (MGS) algorithm, the projection of the observation vector $\mathbf{y}$ onto the space spanned by the columns in the active set is used to form a residual vector needed for the next iteration of the algorithm. When sparsity level k is unknown, the norm residual vector is computed and used as the stopping criteria of OMP. Numerous modifications of OMP with enhanced performance have been proposed in literature. For instance, instead of choosing a single column in each iteration of OMP, StOMP [16] selects and explores all columns having correlation with a residual vector that is greater than a pre-determined threshold. GOMP [17] employs the similar idea, but instead of thresholding, a fixed number of columns is selected per iteration. CoSaMP algorithm [18] identifies columns with largest proximity to the residual vector, uses them to find a least-squares approximation of the unknown signal, and retains only significantly large entries in the resulting approximation. When the unknown signal is a random quantity, rakeness-Based OMP approach [19], attempts to design the measurement matrix by taking into account the second-order statistics of the signal to increase the expected energy of as subset of entries of $\mathbf{y}$. Additionally, necessary and sufficient conditions for exact reconstruction of sparse signals using OMP have been established. Examples of such results include analysis under Restricted Isometry Property (RIP) [20–22], and recovery conditions based on Mutual Incoherence Property (MIP) and Exact Recovery Condition (ERC) [23–25]. For the case of random measurements, performance of OMP was analyzed in [26,27]. Tropp et al. in [26] showed that in the noise-free scenario, $\mathcal{O}(k \log m)$ measurements is adequate to recover k-sparse m-dimensional signals with high probability. In [28], this result was extended to the asymptotic setting of noisy measurements in high signal-to-noise ratio (SNR) under the assumption that the entries of $\mathbf{A}$ are i.i.d. Gaussian and that the length of the unknown vector approaches infinity. Recently, the asymptotic sampling complexity of OMP and GOMP is improved to $\mathcal{O}(k \log \frac{m}{k})$ in [29] and [30], respectively.

Recently, performance of OLS was analyzed in the sparse signal recovery settings with deterministic coefficient matrices. In [31], OLS was analyzed in the noise-free scenario under Exact Recovery Condition (ERC), first introduced in[23]. Herzet et al. [32] provided coherence-based conditions for sparse recovery of signals via OLS when the nonzero components of $\mathbf{x}$ obey certain decay conditions. In [33], sufficient conditions for exact recovery are stated when a subset of true indices is available. In [34] an extension of OLS that employs the idea of [16,17] and identifies multiple indices in each iteration is proposed and its performance is analyzed under RIP. However, all the existing analysis and performance guarantees for OLS pertain to non-random measurements and cannot directly be applied to random coefficient matrices. For instance, the main results in the notable work [29] relies on the assumption of having dictionaries with ℓ_2-norm normalized columns while this obviously does not hold in the scenarios where the coefficient matrix is composed of entries that are drawn from a Gaussian distribution.

3) *Branch-and-bound schemes.* Recently, greedy search heuristics that rely on OMP and OLS to traverse a search tree along paths that represent promising candidates for the support of $\mathbf{x}$ have been proposed. For instance, [35,36] exploit the selection criterion of OMP to construct the search graph while [37,38] rely on OLS to efficiently traverses the search tree. Although these methods empirically improve the performance of greedy algorithms, they are characterized by exponential computational complexity in at least one parameter and hence are prohibitive in applications dealing with high-dimensional signals.

1.1. Contributions

Motivated by the need for fast and accurate sparse recovery in large-scale setting, in this paper we propose a novel algorithm that efficiently exploits recursive relation between components of the optimal solution to the original ℓ_0-constrained least-squares problem (2). The proposed algorithm, referred to as Accelerated Orthogonal Least-Squares (AOLS), similar to GOMP [17] and MOLS [34] exploits the observation that columns having strong correlation with the current residual are likely to have strong correlation with residuals in subsequent iterations; this justifies selection of multiple columns in each iteration and formulation of an overdetermined system of linear equation having solution that is generally more accurate than the one found by OLS or OMP. However, compared to MOLS, our proposed algorithm is orders of magnitude faster and thus more suitable for high-dimensional data applications.

We theoretically analyze the performance of the proposed AOLS algorithm and, by doing so, establish conditions for the exact recovery of the sparse vector $\mathbf{x}$ from measurements $\mathbf{y}$ in (1) when the entries of the coefficient matrix $\mathbf{A}$ are drawn at random from a Gaussian distribution – the first such result under these assumptions for an OLS-based algorithm. We first present conditions which ensure that, in the noise-free scenario, AOLS with high probability recovers the support of $\mathbf{x}$ in k iterations (recall that k denotes the number of non-zero entries of $\mathbf{x}$). Adopting the framework in [26], we further find a lower bound on the probability of performing exact sparse recovery in k iterations and demonstrate that with $\mathcal{O}\left(k \log \frac{m}{k+L-1}\right)$ measurements AOLS succeeds with probability arbitrarily close to one. Moreover, we extend our analysis to the case of noisy measurements and show that similar guarantees hold if the nonzero element of $\mathbf{x}$ with the smallest magnitude satisfies certain condition. This condition implies that to ensure exact support recovery via AOLS in the presence of additive ℓ_2-bounded noise, SNR should scale linearly with sparsity level k. Our procedure for determining requirements that need to hold for AOLS to perform exact reconstruction follows the analysis of OMP in [26,28,27], although with two major differences. First, the variant of OMP analyzed in [26,28,27] implicitly assumes that the columns of $\mathbf{A}$ are ℓ_2-normalized which clearly does not hold if the entries of $\mathbf{A}$ are drawn from a Gaussian distribution. Second, the analysis in [26] is for noiseless measurements while [28,27] essentially assume that SNR is infinite as $k \to \infty$. To the contrary, our analysis makes neither of those two restrictive assumptions. Moreover, we show that if m is sufficiently greater than k, the proposed AOLS algorithm requires $\mathcal{O}\left(k \log \frac{m}{k+L-1}\right)$ random measurements to perform exact recovery in both noiseless and bounded noise scenarios; this is fewer than $\mathcal{O}(k \log(m-k))$ that was found in [28,27] to be the asymptotic sampling complexity for OMP, and $\mathcal{O}\left(k \log \frac{m}{k}\right)$ that was found for MOLS, GOMP, and BP in [34,30,39]. Additionally, our analysis framework is recognizably different from that of [29] for OLS. First, in [29] it is assumed that $\mathbf{A}$ has ℓ_2-normalized columns, and hence the analysis in [29] does not apply to the case of Gaussian matrices, the scenario addressed in this paper for our proposed algorithm. Further, the main result of [29] (see Theorem 3 in [29]) states that OLS exactly recovers a k-sparse vector in at most $6k$ iterations if

Download English Version:

https://daneshyari.com/en/article/6951597

Download Persian Version:

https://daneshyari.com/article/6951597

Daneshyari.com