Appl. Comput. Harmon. Anal. 41 (2016) 412–435

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Hard thresholding pursuit algorithms: Number of iterations ☆

Jean-Luc Bouchot, Simon Foucart *, Pawel Hitczenko

A R T I C L E I N F O

Article history:
Received 13 February 2015
Received in revised form 27 November 2015
Accepted 1 March 2016
Available online 7 March 2016
Communicated by Spec. Issue Guest Editor

MSC:
65F10
65J20
15A29
94A12

Keywords:
Compressive sensing
Uniform sparse recovery
Nonuniform sparse recovery
Random measurements
Iterative algorithms
Hard thresholding

A B S T R A C T

The Hard Thresholding Pursuit algorithm for sparse recovery is revisited using a new theoretical analysis. The main result states that all sparse vectors can be exactly recovered from compressive linear measurements in a number of iterations at most proportional to the sparsity level as soon as the measurement matrix obeys a certain restricted isometry condition. The recovery is also robust to measurement error. The same conclusions are derived for a variation of Hard Thresholding Pursuit, called Graded Hard Thresholding Pursuit, which is a natural companion to Orthogonal Matching Pursuit and runs without a prior estimation of the sparsity level. In addition, for two extreme cases of the vector shape, it is shown that, with high probability on the draw of random measurements, a fixed sparse vector is robustly recovered in a number of iterations precisely equal to the sparsity level. These theoretical findings are experimentally validated, too.

1. Introduction

This paper deals with the standard compressive sensing problem, i.e., the reconstruction of vectors $\mathbf{x} \in \mathbb{C}^N$ from an incomplete set of $m \ll N$ linear measurements organized in the form $\mathbf{y} = \mathbf{A}\mathbf{x} \in \mathbb{C}^m$ for some matrix $\mathbf{A} \in \mathbb{C}^{m \times N}$. It is now well known that if $\mathbf{x}$ is s-sparse (i.e., has only s nonzero entries) and if $\mathbf{A}$ is a random matrix whose number m of rows scales like s times some logarithmic factors, then the reconstruction of $\mathbf{x}$ is achievable via a variety of methods. The ℓ_1-minimization is probably the most popular one, but simple iterative algorithms do provide alternative methods. We consider here the hard thresholding pursuit (HTP) algorithm [6] as well as a novel variation and we focus on the number of iterations needed for the reconstruction. This reconstruction is addressed in two settings:

☆ S.F. and J.-L.B. partially supported by NSF (DMS-1120622), P.H. by Simons Foundation (grant 208766).
* Corresponding author.
E-mail address: simon.foucart@centraliens.net (S. Foucart).

http://dx.doi.org/10.1016/j.acha.2016.03.002

- An idealized situation, where the vectors $\mathbf{x} \in \mathbb{C}^N$ are exactly sparse and where the measurements $\mathbf{y} \in \mathbb{C}^m$ are exactly equal to $\mathbf{Ax}$. In this case, the exact reconstruction of $\mathbf{x} \in \mathbb{C}^N$ is targeted.
- A realistic situation, where the vectors $\mathbf{x} \in \mathbb{C}^N$ are not exactly sparse and where the measurements $\mathbf{y} \in \mathbb{C}^m$ contain errors $\mathbf{e} \in \mathbb{C}^m$, i.e., $\mathbf{y} = \mathbf{Ax} + \mathbf{e}$. In this case, only an approximate reconstruction of $\mathbf{x} \in \mathbb{C}^N$ is targeted. Precisely, the reconstruction error should be controlled by the sparsity defect and by the measurement error. The sparsity defect can be incorporated in the measurement error if $\mathbf{y} = \mathbf{Ax} + \mathbf{e}$ is rewritten as $\mathbf{y} = \mathbf{Ax}_S + \mathbf{e}'$ where S is an index set of s largest absolute entries of $\mathbf{x}$ and $\mathbf{e}' := \mathbf{Ax}_{\overline{S}} + \mathbf{e}$.

We shall mainly state and prove our results in the realistic situation. They specialize to the idealized situation simply by setting $\mathbf{e}' = \mathbf{0}$. In fact, setting $\mathbf{e}' = \mathbf{0}$ inside the proofs would simplify them considerably.

Let us now recall that (HTP) consists in constructing a sequence $(\mathbf{x}^n)$ of s-sparse vectors, starting with an initial s-sparse $\mathbf{x}^0 \in \mathbb{C}^N$ — we take $\mathbf{x}^0 = \mathbf{0}$ — and iterating the scheme[1]

$$S^n := \text{index set of } s \text{ largest absolute entries of } \mathbf{x}^{n-1} + \mathbf{A}^*(\mathbf{y} - \mathbf{Ax}^{n-1}), \tag{HTP$_1$}$$

$$\mathbf{x}^n := \operatorname{argmin}\{\|\mathbf{y} - \mathbf{Az}\|_2, \operatorname{supp}(\mathbf{z}) \subseteq S^n\}, \tag{HTP$_2$}$$

until a stopping criterion is met. It had been shown [10] that, in the idealized situation, exact reconstruction of every s-sparse $\mathbf{x} \in \mathbb{C}^N$ is achieved in s iterations of (HTP) with $\mathbf{y} = \mathbf{Ax}$ provided the coherence of the matrix $\mathbf{A} \in \mathbb{C}^{m \times N}$ satisfies $\mu < 1/(3s)$ (note that this condition can be fulfilled when $m \asymp s^2$). Exact and approximate reconstructions were treated in [6], where it was in particular shown that every s-sparse $\mathbf{x} \in \mathbb{C}^N$ is the limit of the sequence $(\mathbf{x}^n)$ produced by (HTP) with $\mathbf{y} = \mathbf{Ax}$ provided the $3s$th restricted isometry constant of the measurement matrix $\mathbf{A} \in \mathbb{C}^{m \times N}$ obeys $\delta_{3s} < 1/\sqrt{3} \approx 0.577$ (note that this condition can be fulfilled when $m \asymp s\ln(N/s)$). As a reminder, the kth restricted isometry constant δ_k of $\mathbf{A}$ is defined as the smallest constant $\delta \geq 0$ such that

$$(1-\delta)\|\mathbf{z}\|_2^2 \leq \|\mathbf{Az}\|_2^2 \leq (1+\delta)\|\mathbf{z}\|_2^2 \qquad \text{for all } k\text{-sparse } \mathbf{z} \in \mathbb{C}^N.$$

In fact, it was shown in [6] that the convergence is achieved in a finite number $\bar{n}$ of iterations. In the idealized situation, it can be estimated as

$$\bar{n} \leq \left\lceil \frac{\ln(\sqrt{2/3}\|\mathbf{x}\|_2/\xi)}{\ln(1/\rho_{3s})} \right\rceil, \qquad \rho_{3s} := \sqrt{\frac{2\delta_{3s}^2}{1-\delta_{3s}^2}} < 1, \qquad \xi := \min_{j \in \operatorname{supp}(\mathbf{x})} |x_j|. \tag{1}$$

This paper establishes that the number of iterations can be estimated independently of the shape of $\mathbf{x}$: under a restricted isometry condition, it is at most proportional to the sparsity s, see Theorem 5 and a robust version in Theorem 6. This is reminiscent of the work of T. Zhang [15] on orthogonal matching pursuit (OMP), see also [7, Theorem 6.25] where it is proved that $\bar{n} \leq 12s$ provided that $\delta_{13s} < 1/6$.

However, (HTP) presents a significant drawback in that a prior estimation of the sparsity s is required to run the algorithm, while (OMP) does not (although stopping (OMP) at iteration $12s$ does require this estimation). We therefore consider a variation of (HTP) avoiding the prior estimation of s. We call it graded hard thresholding pursuit (GHTP) algorithm, because the index set has a size that increases with the iteration. Precisely, starting with $\mathbf{x}^0 = \mathbf{0}$, a sequence $(\mathbf{x}^n)$ of n-sparse vectors is constructed according to

[1] Exact arithmetic is assumed and, among several candidates for the index set of largest absolute entries, the smallest one in lexicographic order is always chosen.