A Nonlinear Extension of Compressive Sensing

Henrik Ohlsson, Allen Y. Yang, Roy Dong, and S. Shankar Sastry
Dept. of Electrical Engineering & Computer Sciences, University of California, Berkeley, USA
Email: ohlsson,yang,roydong,sastry@eecs.berkeley.edu

Abstract—In many areas where compressive sensing is being used today, the relationship between the measurements and the unknowns could be nonlinear. Traditional treatment of such nonlinear relationships has been to approximate the nonlinearity via a linear model and the subsequent unmodeled dynamics as noise. The ability to more accurately characterize nonlinear models has the potential to improve the results in both existing compressive sensing applications and those where a linear approximation does not suffice, e.g., phase retrieval. In this study, we extend the classical compressive sensing framework to a second-order Taylor expansion of the nonlinearity. Using a lifting technique, we show that the sparse signal can be recovered exactly when the sampling rate is sufficiently high. We further derive efficient numerical algorithms to recover sparse signals in second-order nonlinear systems via alternating direction method of multipliers (ADMM).

I. INTRODUCTION

It has recently been shown that compressive sensing (CS) can be extended to nonlinear models. More specifically, the relatively new topic of nonlinear compressive sensing (NLCS) studies the general problem of finding the sparsest signal $x$ satisfying:

$$y_i = f_i(x), \quad y_i \in \mathbb{R}, \ i = 1, \ldots, N,$$

where each $f_i : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function. Compared to CS, the literature on NLCS is still very limited. The interested reader is referred to [1], [2], [3] and references therein.

In this paper, we will restrict our attention from rather general nonlinear systems to a quadratic compressive sensing problem. More specifically, we consider the following problem:

$$\min_{x \in \mathbb{R}^n} \|x\|_0 \quad \text{subj. to} \quad y_i = a_i^T x + b_i^T x + x^T Q_i x, \ i = 1, \ldots, N.$$  (1)

where $a_i \in \mathbb{R}$, $b_i \in \mathbb{R}^n$, $y_i \in \mathbb{R}$, and $Q_i \in \mathbb{R}^{n \times n}$, $i = 1, \ldots, N$. In a sense, extending CS to quadratic models makes it possible to apply the principles of CS to a second-order Taylor expansion of the nonlinear relationship in (1), while traditional CS mainly considers its first-order Taylor expansion.

II. CONVEX RELAXATION VIA LIFTING

As optimizing the $\ell_0$-norm function in (2) is not efficient and known to be a non-convex combinatorial problem, we first introduce a convex relaxation of the problem.

Define $\Phi_i = \begin{bmatrix} a_i \\ b_i^T / 2 \\ Q_i \end{bmatrix}$ and $X = \begin{bmatrix} x \\ x^T \end{bmatrix}$, both matrices of dimensions $(n + 1) \times (n + 1)$. Then (2) is equivalent to

$$\min_X \|X\|_0 \quad \text{subj. to} \quad y_i = \text{Tr}(\Phi_i X), \ i = 1, \ldots, N,$$

where $\text{rank}(X) = 1$, $X_{1,1} = 1$, $X \succeq 0$. (3c)

where the zero norm counts the number of nonzero elements in the matrix $X$. This problem is still non-convex and combinatorial. Inspired by recent literature on matrix completion and sparse PCA, we relax the problem into the following convex semidefinite program (SDP):

$$\min_X \text{Tr}(X) + \lambda \|X\|_1$$

subject to

$$y_i = \text{Tr}(\Phi_i X), \ i = 1, \ldots, N,$$

$$X_{1,1} = 1, X \succeq 0.$$  (4c)

where we define $\|X\|_1$ as the absolute sum of all elements of $X$ and $\lambda \geq 0$ is a design parameter. We refer to the approach as quadratic basis pursuit (QBP). See [4] for details.

III. THEORETICAL ANALYSIS

It is convenient to introduce the linear operator $B$ as

$$B : X \in \mathbb{R}^{n \times n} \rightarrow \{\text{Tr}(\Phi_i X)\}_{i \leq i \leq N} \in \mathbb{R}^N,$$

and $y$ as the vector of stacked measurements, $y = [y_1 \ y_2 \ \ldots \ y_N]$. We further need a generalization of the restricted isometry property (RIP) and Mutual Coherence.

Definition 1 (RIP). A linear operator $B(\cdot)$ as defined in (5) is $(\epsilon, k)$-RIP if $\|B(x)\|_2^2 - 1 < \epsilon$ for all $\|x\|_0 \leq k$ and $X \neq 0$.

Definition 2 (Mutual Coherence). For a matrix $A$, define the mutual coherence as $\mu(A) = \max_{i \leq j \leq n, i \neq j} |A_{ij}| / \|A\|_{2}$. By an abuse of notation, let $B$ be the matrix satisfying $y = BX^*$ with $X^*$ being the vectorized version of $X$.

Theorem 1 (Guaranteed recovery using RIP and Mutual Coherence). Let $k$ be the sparest solution to (2). The solution of QBP $\hat{X}$ is equal to

$$\begin{bmatrix} 1 \\ \hat{x}^T \end{bmatrix} \text{ if it has rank } 1 \text{ and } B(\cdot) \text{ is } (\epsilon, 2)(\hat{X})_0 \text{-RIP with }$$

$$\epsilon < 1 \text{ or } \|\hat{X}\|_0 < 0.5(1 + \mu(B)).$$

We will also discuss several other interesting theoretical results and how to handle noise via a stable extension. Finally, we have presented an efficient and stable ADMM-based numerical solver in [5].

ACKNOWLEDGMENT

Ohlsson is partially supported by the Swedish Research Council in the Linnaeus center CADICS, the European Research Council under the advanced grant LEARN, contract 267381, by a postdoctoral grant from the Linnaeus center CADICS, the European Research Council under the advanced grant LEARN, contract 267381, by a postdoctoral grant from the Sweden-America Foundation, donated by ASEA’s Fellowship Fund, and by a postdoctoral grant from the Swedish Research Council. Yang is supported in part by ARO 63092-MA-II and DARPA FA8650-11-1-7153.

REFERENCES


