Signal Processing 130 (2017) 57–63

Contents lists available at ScienceDirect

Signal Processing

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

Short communication

Fast reconstruction algorithm for perturbed compressive sensing based on total least-squares and proximal splitting

Reza Arablouei

Commonwealth Scientific and Industrial Research Organisation (CSIRO), Pullenvale, QLD, Australia

ARTICLE INFO

Article history:
Received 9 March 2016
Received in revised form
24 May 2016
Accepted 10 June 2016
Available online 11 June 2016

Keywords:
Non-convex optimization
Perturbed compressive sensing
Proximal-gradient method
Rayleigh quotient
Sparse reconstruction
Sparse total least-squares

ABSTRACT

We consider the problem of finding a sparse solution for an underdetermined linear system of equations when the known parameters on both sides of the system are subject to perturbation. This problem is particularly relevant to reconstruction in fully-perturbed compressive-sensing setups where both the projected measurements of an unknown sparse vector and the knowledge of the associated projection matrix are perturbed due to noise, error, mismatch, etc. We propose a new iterative algorithm for tackling this problem. The proposed algorithm utilizes the proximal-gradient method to find a sparse total least-squares solution by minimizing an ℓ_1-regularized Rayleigh-quotient cost function. We determine the step-size of the algorithm at each iteration using an adaptive rule accompanied by back-tracking line search to improve the algorithm's convergence speed and preserve its stability. The proposed algorithm is considerably faster than a popular previously-proposed algorithm, which employs the alternating-direction method and coordinate-descent iterations, as it requires significantly fewer computations to deliver the same accuracy. We demonstrate the effectiveness of the proposed algorithm via simulation results.

1. Introduction

The theory of compressive sensing states that an N-dimensional vector that has only $K \ll N$ nonzero entries can be recovered from its $M < N$ projections when M is sufficiently set and the projection matrix has certain properties. The projection matrix is generally the product of a sensing matrix and a sparse representation matrix. Conventionally, when reconstructing the unknown sparse vector, the projection matrix is assumed to be perfectly known while the projections may be subject to perturbation stemming from background noise or measurement error [1–3]. In many applications, the perfect knowledge of the projection matrix is infeasible and only a perturbed version of it is available for the reconstruction of the original sparse vector. Examples of such applications are grid-based approaches to time-delay/Doppler-shift/direction-of-arrival/position estimation in communications/radar systems or spectrum sensing in cognitive radio networks [4–10], mobile electrocardiogram (ECG) monitoring [11], X-ray imaging [12], plant biomass characterization [13], hyperspectral unmixing [14], information security [15,16], and high-dimensional linear regression [17].

The presence of perturbation in both the projection matrix and the vector of projections has given rise to the so-called perturbed compressive sensing (PCS) paradigm. The reconstruction of the target sparse vector in PCS amounts to solving a fully-perturbed underdetermined system of linear equations (SLE) under a sparsity assumption. The effects of perturbation on the projection matrix as well as the vector of projections have been analysed and relevant theoretical performance bounds have been reported in several works including [18–21].

The total least-squares (TLS) method is an effective way of solving a fully-perturbed SLE, albeit disregarding any possible sparseness in the target solution vector. The TLS is a linear fitting technique that accounts for perturbations on both sides of an SLE [22–26]. An augmented matrix can be formed by concatenating the perturbed parameter matrix of the left-hand side of an SLE with the perturbed parameter vector of its right-hand side. The optimal TLS solution to the SLE is related to the singular vector of this augmented matrix that corresponds to its smallest singular value [24]. This singular value is the minimum value for the Rayleigh-quotient of the covariance of the augmented matrix [25–28]. Besides, it is appreciated that adding an ℓ_1-norm penalty as a regularization term to a least-squares cost function promotes sparsity in its minimizer [29]. Therefore, one can expect to attain a sparse solution to a fully-perturbed underdetermined SLE by combining the concepts of TLS fitting and ℓ_1-norm regularization. In [4], two algorithms based on ℓ_1-regularized TLS estimation are proposed, one of which is near-optimal and the other is sub-optimal but with reduced complexity.

E-mail address: reza.arablouei@csiro.au

http://dx.doi.org/10.1016/j.sigpro.2016.06.009
0165-1684/

The matrix-uncertainty generalized approximate message-passing (MU-GAMP) algorithm, proposed in [30], takes a Bayesian estimation approach to solve the PCS recovery problem. It is based on the generalized approximate message-passing (GAMP) algorithm [31,32] and exploits the prior knowledge about the probability distribution of the target sparse vector. Another popular approach for reconstruction in PCS is to employ a greedy strategy to enforce sparsity on the TLS (or any other perturbation-compensated) estimate. Some of the greedy PCS reconstruction algorithms combine the TLS estimation with a greedy support-detection method; others modify the classical greedy compressive-sensing reconstruction algorithms, such as the orthogonal matching pursuit algorithm, so that they can account for the perturbation in the projection matrix as well as the perturbation in the vector of projections. The algorithms proposed in [6,11,33–35] are a few examples. These algorithms along with those based on the ℓ_1-regularized TLS estimation are among the most computationally efficient reconstruction algorithms for PCS.

In this paper, we propose a new reconstruction algorithm for PCS that finds a sparse total least-squares estimate by minimizing an ℓ_1-regularized Rayleigh-quotient cost function. To this end, we utilize the proximal-gradient method, which is also known as the forward-backward splitting method, [36–38] and determine its step-sizes through an adaptive scheme accompanied by backtracking line search. The proximal-gradient method is suitable for minimizing composite cost functions comprised of two additive terms, one of which is differentiable and the other is convex and admits a proximity operator [39]. It is a two-stage iterative algorithm that addresses each term in the composite cost function separately. At each iteration, it moves the estimate along the opposite direction of the gradient of the differentiable term (forward gradient-descent step), then adjusts the estimate by applying the proximity operator of the other term (backward gradient-descent step) [36]. The proximal-gradient method is a simple practical algorithm that can often be implemented with relatively low complexity.

We assume that the perturbations have Gaussian distribution. The case of perturbations with Poisson distribution has been studied in [12]. Thanks to the computational efficiency of the proximal-gradient method, the computational complexity of the proposed algorithm is of order $O(NK)$ to $O(N^2)$ at each iteration, depending on the sparseness of the estimates. Consequently, the proposed algorithm demands significantly fewer computations than its closest contender, the suboptimal algorithm of [4], which has a per-iteration computational complexity order of $O(NMK)$ to $O(N^2M)$. Notably, we achieve this improvement in complexity with no sacrifice of estimation accuracy.

We provide simulation examples to examine the performance of the proposed algorithm in comparison with the algorithm of [4]. The simulation results corroborate the efficacy of the proposed algorithm.

2. Problem

We consider the problem of finding an estimate of the vector $\mathbf{x}_o \in \mathbb{R}^{N\times 1}$ as the solution of the following underdetermined system of linear equations:

$$\mathbf{A}_o \mathbf{x}_o = \mathbf{b}_o. \tag{1}$$

The matrix $\mathbf{A}_o \in \mathbb{R}^{M\times N}$ has fewer rows than columns, i.e., $M < N$, and the target vector $\mathbf{x}_o$ is sparse with $K < M$ nonzero entries. We do not observe $\mathbf{A}_o$ and $\mathbf{b}_o \in \mathbb{R}^{M\times 1}$ directly. Instead, we have access to their perturbed versions $\mathbf{A}$ and $\mathbf{b}$, respectively, which are given by

$$\mathbf{A} = \mathbf{A}_o - \mathbf{E}_o \text{ and } \mathbf{b} = \mathbf{b}_o - \mathbf{e}_o \tag{2}$$

where $\mathbf{E}_o \in \mathbb{R}^{M\times N}$ and $\mathbf{e}_o \in \mathbb{R}^{M\times 1}$ are unknown perturbations. The source of perturbation can be background noise, measurement error, quantization error, coarse discretization of the parameter space, sampling jitter, mismatch between the true values and the existing knowledge about them, etc. Substituting (2) into (1) gives the following errors-in-variables equation:

$$(\mathbf{A} + \mathbf{E}_o)\mathbf{x}_o = (\mathbf{b} + \mathbf{e}_o).$$

In the context of compressive sensing, $\mathbf{A}_o$ is known as the unperturbed projection or measurement matrix and $\mathbf{b}_o$ is called the vector of unperturbed projections or measurements. The parameters with index o are unobservable (hidden). We are primarily interested in estimating (recovering) the sparse target vector $\mathbf{x}_o$ from the known but perturbed parameters $\mathbf{A}$ and $\mathbf{b}$. As shown in [4], this problem, i.e., reconstruction in PCS, can be cast as an ℓ_1-regularized total least-squares problem where estimates of $\mathbf{x}_o$, $\mathbf{e}_o$, and $\mathbf{E}_o$ are found by solving

$$\min_{\mathbf{x},\mathbf{e},\mathbf{E}} \left(\|\mathbf{e}\|_2^2 + \|\mathbf{E}\|_F^2 + \lambda\|\mathbf{x}\|_1 \right)$$
$$\text{subject to } (\mathbf{A} + \mathbf{E})\mathbf{x} = \mathbf{b} + \mathbf{e}. \tag{3}$$

Here, $\lambda > 0$ is the regularization parameter while $\|\cdot\|_2$, $\|\cdot\|_F$, and $\|\cdot\|_1$ stand for ℓ_2, Frobenius, and ℓ_1 norms, respectively. The estimates for $\mathbf{x}_o$, $\mathbf{e}_o$, and $\mathbf{E}_o$ produced by solving (3) are optimal in the maximum *a posteriori* (MAP) sense when the entries of $\mathbf{x}_o$ are independently drawn from a zero-mean common Laplace distribution with parameter $2/\lambda$ and the entries of $\mathbf{e}_o$ and $\mathbf{E}_o$ are independent identically-distributed Gaussian with zero mean and equal variance. According to Lemma 1 of [4], the constrained optimization problem (3) is equivalent to two unconstrained optimization problems, one involving the variables $\mathbf{x}$ and $\mathbf{E}$ as

$$\min_{\mathbf{x},\mathbf{E}} \left[\left\| (\mathbf{A} + \mathbf{E})\mathbf{x} - \mathbf{b} \right\|_2^2 + \|\mathbf{E}\|_F^2 + \lambda\|\mathbf{x}\|_1 \right] \tag{4}$$

and the other involving only $\mathbf{x}$ as

$$\min_{\mathbf{x}} \left(\frac{\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2}{\|\mathbf{x}\|_2^2 + 1} + \lambda\|\mathbf{x}\|_1 \right). \tag{5}$$

In [4], two iterative algorithms are proposed for solving the ℓ_1-regularized total least-squares problem. The first algorithm solves an equivalent form of (5) expressed as

$$\min_{\mathbf{x}} \frac{\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2}{\|\mathbf{x}\|_2^2 + 1}$$
$$\text{subject to } \|\mathbf{x}\|_1 \le \delta \tag{6}$$

where δ is an estimate of the ℓ_1-norm of the optimal solution evaluated through a cross-validation scheme. This algorithm has inner and outer iteration loops that are based on a variant of the branch-and-bound method [40] and the bisection method [41], respectively. Although this algorithm is guaranteed to converge to an arbitrarily small neighbourhood of the global solution of (6), it is computationally expensive as its complexity is not necessarily of polynomial order.

The second algorithm proposed in [4] solves (4) using an alternating-direction approach. It successively alternates between two steps: 1) estimating $\mathbf{x}_o$ given the last estimate of $\mathbf{E}_o$ using coordinate-descent iterations and 2) estimating $\mathbf{E}_o$ given the last estimate of $\mathbf{x}_o$ by solving a straightforward quadratic subproblem. This algorithm, which we will refer to as the alternating-direction coordinate-descent (AD-CD) algorithm, is only guaranteed to converge to a local minimum of the cost function in (4) and not necessarily to its global minimum. However, it is computationally