Journal of Computational Mathematics
Vol.28, No.2, 2010, 170–194.

http://www.global-sci.org/jcm
doi:10.4208/jcm.2009.10-m1007

FIXED-POINT CONTINUATION APPLIED TO COMPRESSED SENSING: IMPLEMENTATION AND NUMERICAL EXPERIMENTS*

Elaine T. Hale   Wotao Yin and Yin Zhang
Department of Computational and Applied Mathematics
Rice University, Houston, Texas, 77005, USA
Email: Elaine.Hale@nrel.gov, wotao.yin@rice.edu, zhang@caam.rice.edu

Abstract

Fixed-point continuation (FPC) is an approach, based on operator-splitting and continuation, for solving minimization problems with ℓ_1-regularization:

$$\min \|x\|_1 + \bar{\mu} f(x).$$

We investigate the application of this algorithm to compressed sensing signal recovery, in which $f(x) = \frac{1}{2}\|Ax - b\|_M^2$, $A \in \mathbb{R}^{m \times n}$ and $m \leq n$. In particular, we extend the original algorithm to obtain better practical results, derive appropriate choices for M and $\bar{\mu}$ under a given measurement model, and present numerical results for a variety of compressed sensing problems. The numerical results show that the performance of our algorithm compares favorably with that of several recently proposed algorithms.

Mathematics subject classification: 65K05, 90C06, 90C25, 90C90
Key words: ℓ_1 regularization, Fixed-point algorithm, Continuation, Compressed sensing, Numerical experiments.

1. Introduction

The fixed-point continuation (FPC) algorithm proposed in [40] can be used to compute sparse solutions for under-determined linear systems $Ax = b$ using the weighted least-squares formulation

$$\min_{x \in \mathbb{R}^n} \|x\|_1 + \frac{\bar{\mu}}{2}\|Ax - b\|_M^2, \tag{1.1}$$

where $A \in \mathbb{R}^{m \times n}$, $m < n$, $\|p\|_M^2 = p^T M p$ and M is positive definite. This paper describes implementation details and usage guidelines for this setting, and summarizes a series of numerical experiments. The experiments simulate compressed sensing applications and provide for direct comparison of FPC with three other state-of-the-art compressed sensing recovery algorithms: GPSR [36], $l1\_ls$ [42], and StOMP [27].

1.1. Background

In some applications, sparse solutions, that is, vectors that contain many zero elements, are preferred over dense solutions that are otherwise equally suitable. This was recognized early in geophysics, where sparse spike train signals are often of interest and data may include large sparse errors [18, 46, 60, 63]. The signal processing community seeks sparse vectors so as to

* Received March 18, 2009 / Revised version received May 24, 2009 / Accepted June 6, 2009 / Published online December 21, 2009 /

describe a signal with just a few waveforms; similarly, statisticians often want to identify a parsimonious set of explanatory variables [17, 28, 50, 51, 64].

A direct, but computationally intractable, method for finding the sparsest solution to an under-determined linear system is to minimize the so-called "ℓ_0-norm", that is, the number of nonzeros in a vector. On the other hand, minimizing or bounding $\|x\|_1$ has long been recognized as a practical substitute, as the two are equivalent under suitable conditions. This yields the so-called basis pursuit problem [17]

$$\min_{x\in\mathbb{R}^n} \{\|x\|_1 \mid Ax = b\}. \tag{1.2}$$

If the "observation" b is contaminated with noise ϵ, i.e.,

$$b = Ax + \epsilon,$$

then an appropriate norm of the residual $Ax - b$ should be minimized or constrained. Such considerations yield a family of related optimization problems. For instance, if the noise is Gaussian then the ℓ_1-regularized least squares problem

$$\min_{x\in\mathbb{R}^n} \|x\|_1 + \frac{\bar{\mu}}{2}\|Ax - b\|_2^2, \tag{1.3}$$

would be appropriate, as would the Lasso problem [64]

$$\min_{x\in\mathbb{R}^n} \left\{\|Ax - b\|_2^2 \mid \|x\|_1 \le t\right\}, \tag{1.4}$$

which is equivalent to (1.3) given appropriate constants $\bar{\mu}$ and t. Note that formulations (1.1) and (1.3) are equivalent since a weighting matrix M can be incorporated in (1.3) by multiplying A and b on the left by $M^{1/2}$. We use the explicitly weighted formulation (1.1) because it arises naturally from the stochastic measurement model introduced in Section 3.1.

1.2. l_1-Regularization and Compressed Sensing

Compressed Sensing is the name assigned to the idea of encoding a large sparse signal using a relatively small number of linear measurements, and decoding the signal either through minimizing the ℓ_1-norm (or its variants) or employing a greedy algorithm. The current burst of research in this area is traceable to new results reported by Candes *et al.* [12–14], Donoho *et al.* [25, 26, 68] and others [59, 65]. Applications of compressed sensing include compressive imaging [62, 73, 74], medical imaging [48], multi-sensor and distributed compressed sensing [3], analog-to-information conversion [43–45, 67], and missing data recovery [81]. Compressed sensing is attractive for these and other potential applications because one can obtain a given quantity of information with fewer measurements in exchange for some additional post-processing.

In brief, compressed sensing theory shows that a sparse signal of length n can be recovered from $m < n$ measurements by solving an appropriate variant of (1.2), (1.3), (1.4), etc., provided that the $m \times n$ measurement matrix A possesses certain "good" properties. To date, random matrices and matrices whose rows are taken from certain orthonormal matrices have been proven to be "good". These matrices are invariably large and dense, which contradicts the usual assumption of optimization solvers that large-scale problems appear with sparse data. The size and density of the data involved further suggest that solution algorithms should not require large linear system solves or matrix factorizations, and should take full advantage of available fast transforms like FFT and DCT. Thus it is necessary to develop dedicated algorithms for compressed sensing signal reconstruction that have the aforementioned properties and are as fast and memory-efficient as possible.