Applied Numerical Mathematics 60 (2010) 452–472

Contents lists available at ScienceDirect

Applied Numerical Mathematics

www.elsevier.com/locate/apnum

Sparse approximate solution of partial differential equations ☆

Sadegh Jokar [a], Volker Mehrmann [a,*], Marc E.Pfetsch [b], Harry Yserentant [a]

[a] *Institut für Mathematik, TU Berlin, Str. des 17. Juni 136, 10623 Berlin, Germany*
[b] *Institute for Mathematical Optimization, TU Braunschweig, Pockelsstr. 14, 38106 Braunschweig, Germany*

ARTICLE INFO

Article history:
Received 21 February 2009
Received in revised form 4 October 2009
Accepted 6 October 2009
Available online 14 October 2009

MSC:
65N50
65K05
65F20
65F50

Keywords:
Partial differential equation
Sparse solution
Dictionary
Compressed sensing
Restricted isometry property
Mutual incoherence
Hierarchical basis
Linear programming

ABSTRACT

A new concept is introduced for the adaptive finite element discretization of partial differential equations that have a sparsely representable solution. Motivated by recent work on compressed sensing, a recursive mesh refinement procedure is presented that uses linear programming to find a good approximation to the sparse solution on a given refinement level. Then only those parts of the mesh are refined that belong to nonzero expansion coefficients. Error estimates for this procedure are refined and the behavior of the procedure is demonstrated via some simple elliptic model problems.

1. Introduction

The sparse representation of functions via a linear combination of a small number of basic functions has recently received a lot of attention in several mathematical fields such as approximation theory [19,36,41,42] as well as signal and image processing [7,9–13,15,22–28]. In terms of representations of functions, we can describe the problem as follows. Consider a linearly dependent set of n functions ϕ_i, $i = 1, 2, \dots, n$ (*a dictionary* [16]) and a function f represented as

$$f = \sum_{i=1}^{n} x_i \phi_i.$$

Since the set of functions is not linearly independent, this representation is not unique and we may want to determine the sparsest representation, i.e., a representation with a maximal number of vanishing coefficients among $x_1, \dots, x_n$. In the setting of numerical linear algebra, this problem can be formulated as follows. Consider a linear system

☆ Supported by the *Deutsche Forschungsgemeinschaft* through the DFG Research Center Matheon *Mathematics for key technologies* in Berlin.

* Corresponding author.
E-mail addresses: jokar@math.tu-berlin.de (S. Jokar), mehrmann@math.tu-berlin.de (V. Mehrmann), m.pfetsch@tu-bs.de (M.E. Pfetsch), yserenta@math.tu-berlin.de (H. Yserentant).

0168-9274/$30.00 © 2009 IMACS. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.apnum.2009.10.003

$$\Phi x = b, \tag{1}$$

with $\Phi \in \mathbb{R}^{m,n}$, where $m \leqslant n$ and $b \in \mathbb{R}^m$. The columns of the matrix Φ and the right-hand side b represent the functions ϕ_i and the function f, respectively, with respect to some basis of the relevant function space. The problem is then to find the sparsest possible solution x, i.e., x has as many zero components as possible. This optimization problem is in general NP-hard [30,37]. Starting from the work of [15], however, a still growing number of articles have developed sufficient conditions that guarantee that an (approximate) sparse solution $\hat{x}$ to (1) can be obtained by solving the linear program

$$\min \|x\|_1, \quad \text{s.t.} \quad \Phi x = b \quad \big(\text{or } \|\Phi x - b\| \leqslant \varepsilon\big),$$

which can be done in polynomial time, see [34,35] and [33] for a discussion. We will give a brief survey of this theory in Section 2.2.

In the literature, the development has mostly focused on the construction of appropriate coding matrices Φ that allow for the sparse representation of a large class of functions (signals or images). Furthermore, properties of the columns of the matrix (or the dictionary) have been investigated, which guarantee that the computation of the sparse solution can be done efficiently via a linear programming approach, see, for instance, [13,24] and [33] with its references. Often the term *compressed sensing* is used for this approach.

In this paper we consider a related but different problem. We are interested in the numerical solution of partial differential equations

$$Lu = f,$$

with a differential operator L, to be solved in a domain $\Omega \subset \mathbb{R}^d$ with smooth boundary Γ and appropriate boundary conditions given on Γ.

Considering a classical Galerkin or Petrov–Galerkin finite element approach, see e.g. [5], one seeks a solution u in some function space $\mathbb{U}$ (which is spanned by $\phi_1, \ldots, \phi_n$), represented as

$$u = \sum_{i=1}^{n} u_i \phi_i. \tag{2}$$

Again we are interested in sparse representations with a maximal number of vanishing coefficients u_i. In contrast to the cases discussed before, here we would like to construct the space $\mathbb{U}$ and the basis functions ϕ_i in the finite element discretization in such a way that first of all a sparse representation of the solution to (2) exists and second that it can be determined efficiently. Furthermore, it would be ideal if the functions ϕ_i could be constructed in a multilevel or adaptive way.

The usual approach to achieve this goal is to use local a posteriori error estimation to determine where a refinement, i.e., the addition of further basis functions is necessary. For example, in the dual weighted residual approach [3] this is done by solving an optimization problem for the error.

In this article, we examine the possibility to use similar approaches as those used in compressed sensing, i.e., to use ℓ_1-minimization and linear programming to perform the adaptive refinement in the finite element method in such a way that the solution is sparsely represented by a linear combination of basis functions. In order to achieve this goal, we propose the following framework.

We determine $u \in \mathbb{U}$ as the solution of the weak formulation

$$(v, Lu - f) = 0 \quad \text{for all } v \in \mathbb{V}.$$

Here, $\mathbb{V}$ is a space of test functions and $(\cdot,\cdot)$ is an appropriate inner product. In the simplest version of a two-level approach, we construct finite-dimensional spaces of coarse and fine basis functions $\mathbb{U}_1^n \subset \mathbb{U}_1^N \subset \mathbb{U}$ and corresponding spaces for coarse and fine test functions $\mathbb{V}_1^n \subset \mathbb{V}_1^N \subset \mathbb{V}$. Then we determine the sparsest solution in $\mathbb{U}_1^N$, such that

$$(v, Lu - f) = 0 \quad \text{for all } v \in \mathbb{V}_1^N \setminus \mathbb{V}_1^n$$

via the solution of an underdetermined system of the form (1). Based on the sparse solution, we determine new coarse and fine spaces $\mathbb{U}_2^n \subset \mathbb{U}_2^N \subset \mathbb{U}$, $\mathbb{V}_2^n \subset \mathbb{V}_2^N \subset \mathbb{V}$, and iterate this procedure (see Section 3.2).

This framework combines the ideas developed in compressed sensing with well-known concepts arising in adaptive and multilevel finite element methods [18]. But instead of using local and global error estimates to obtain error indicators by which the grid refinement is controlled, here the solution of the ℓ_1-minimization problem is used to control the grid refinement and adaptivity.

Many issues of this approach have, however, not yet been resolved, in particular, the theoretical analysis of this approach (see Section 4). We see the following potential advantages and disadvantages of this framework. On the positive side, the ℓ_1-minimization approach allows for an easy automation. We will demonstrate this with some numerical examples in Section 5. On the downside, the analysis of the approach seems to be hard even for classical elliptic problems, see Section 4 and due to the potentially high complexity of the linear programming methods this approach will only be successful, if the procedure needs only a few levels and a small sparse representation of the solution exists, see Section 5.