Available online at www.sciencedirect.com
ScienceDirect
Mathematics and Computers in Simulation 129 (2016) 50–68
MATHEMATICS AND COMPUTERS IN SIMULATION
www.elsevier.com/locate/matcom

Original articles

Data mining and probabilistic models for error estimate analysis of finite element method

Joël Chaskalovic [a,*], Franck Assous [b]

[a] D'Alembert, University Pierre and Marie Curie, Paris, France
[b] Ariel University, 40700 Ariel, Israël

Received 26 November 2014; received in revised form 29 December 2015; accepted 21 March 2016
Available online 24 May 2016

Abstract

In this paper, we propose a new approach based on data mining techniques and probabilistic models to compare and analyze finite element results of partial differential equations. We focus on the numerical errors produced by linear and quadratic finite element approximations. We first show how error estimates contain a kind of numerical uncertainty in their evaluation, which may influence and even damage the precision of finite element numerical results. A model problem, derived from an elliptic approximate Vlasov–Maxwell system, is then introduced. We define some variables as physical predictors, and we characterize how they influence the odds of the linear and quadratic finite elements to be locally "same order" accurate. Beyond this example, this approach proposes a method to compare, between several approximation methods, the accuracy of numerical results.

Keywords: Data mining; Probabilistic models; Error estimate; Finite element

1. Introduction

In many physical problems, we are limited first by our ability to measure observations and then to define the gap between the real problem and the mathematical model used to describe it. This is often referred as the modeling error. In previous papers [1,2], we have proposed to apply data mining techniques to evaluate this error. We relied on the fact that data mining techniques have already proved to be efficient in other contexts which deal with massive data, like in biology [23], medicine [28,27], marketing [25], advertising and communications [9,11].

When we attempt to simulate the problem numerically, we must identify the possible limitations of the numerical techniques employed. This is sometimes referred as the approximation error. In this article, using data mining techniques and probabilistic models, we extend our approach to compare numerical errors produced by different finite element approximations. Indeed, there is a need to investigate ways to quantify uncertainty and its impact, also in the case of numerical approximations, albeit supported by a mathematical theory, i.e. error estimate analysis.

Following [30] or [16], we distinguish between *errors* and *uncertainty* as follows: *errors* are detectable insufficiencies not due to lack of knowledge, whereas *uncertainties* are linked with lack of knowledge. As a consequence,

* Corresponding author.
E-mail address: jch1826@gmail.com (J. Chaskalovic).

http://dx.doi.org/10.1016/j.matcom.2016.03.013

one can consider errors as deterministic quantities, whereas uncertainties are inherently stochastic. This justified the probabilistic framework we will use in this article.

In that spirit, numerical simulations can be subject to uncertainty for instance in boundary conditions, geometry of the physical domain, etc., but they are also subject to uncertainty in their accuracy estimation. Indeed, an error estimate, even accurate, is an approximation to the actual unknown error, due to the presence of uncertainty that we will highlight in the following sections. Hence, uncertainty quantification of a given numerical method can be an interesting step towards its certification, in addition to the well known error estimate analysis.

For this purpose, we consider finite elements, and we aim at identifying a kind of stochastic behavior in finite element error estimates, which justify our stochastic approach later on. Accordingly, large databases that contain numerical approximations will be explored, in order to see whether and where precise finite elements are needed to guarantee accuracy. We will also show the limits of the approach, mainly due to the lack of the reference solution.

To illustrate our method, we introduce a model problem – a quasi-static Vlasov–Maxwell model – which models charged particle beams in plasma physics problems. We derive two numerical approximation methods, based on linear and quadratic finite elements [10], denoted respectively P_1 and P_2. Then, we compare the accuracy between the two implemented methods by modeling the dependency of the odds, which characterizes when P_1 and P_2 finite elements produce equivalent results.

More specifically, our objective is to propose data mining and probabilistic tools to ascertain situations where P_1 and P_2 finite elements lead to equivalent results. Our method will mine stored data of computed approximations to identify the predictors responsible of such a situation. However, at this point, we do not seek under what circumstances one leaves the P_2 finite element for the P_1 ones, to get accurate results for a given concrete application.

This article is organized as follows. In Section 2, we consider the global framework of a general elliptic variational formulation, and we highlight where and how a quantitative uncertainty appears in finite elements error estimates. Then, in Section 3, we introduce a model problem to illustrate our approach, namely a quasi-static paraxial Vlasov–Maxwell approximation. Numerical solutions are then computed by a P_1 and P_2 finite element Particle-In-Cell method. Section 4 will be devoted to the data mining and probabilistic models. After a brief presentation of the data mining tools involved, we will use them on our model problem approximations. Conclusions regarding equivalent numerical P_1 and P_2 approximations will be drawn.

2. Uncertainty in finite elements error estimate

We propose an approach based on statistical and probabilistic models to compare and analyze finite element results of partial differential equations. More precisely, we aim to explore large datasets in order to see whether high-order finite elements are required to guarantee accuracy. Our purpose in this section is to define "the same order" notion. Starting from finite element error analysis, we illustrate how two different finite element approximations can be of the "same order", due to a quantitative uncertainty appearing in the error estimates.

To begin with, let us recall some familiar notions regarding finite element error estimate of partial differential equations. Our main focus will be on a comparison of numerical errors produced by P_1 and P_2 finite element approximations (for more details see [3]). The aim is to show a kind of stochastic behavior in finite element error estimates, justifying a stochastic approach later on.

Since our study focuses mainly on space-dependent part of the model (not the time-dependent one, if any), we consider in what follows an elliptic standard problem, that can be viewed as the stationary problem associated with the time-dependent one. Indeed, when looking at a time-dependent problem, after time discretization, one generally solves a sequence of stationary problems, one for each time step. Hence, for a time-dependent problem, the method proposed here, derived at each time step, can be then accumulated over time.

Let V be a Hilbert space (with norm $\|.\|_V$). Throughout this section, $a(\cdot,\cdot)$ denotes a bilinear, continuous and V-elliptic form defined on $V \times V$, and $L(\cdot)$ a linear continuous form defined on V. We introduce the abstract elliptic variational formulation

$$\begin{cases} \text{Find } u \in V \text{ such that} \\ a(u,v) = L(v), \quad \forall v \in V. \end{cases} \tag{1}$$

Existence and uniqueness of a solution u of (1) is guaranteed by the Lax–Milgram theorem [12]. Let us introduce a finite dimension subset V_h of V, and consider the approximate solution u_h of u, that solves the approximate variational