Journal of Computational Physics 267 (2014) 92–111

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

A weighted ℓ_1-minimization approach for sparse polynomial chaos expansions

Ji Peng [a], Jerrad Hampton [b], Alireza Doostan [b,*]

[a] Mechanical Engineering Department, University of Colorado, Boulder, CO 80309, USA
[b] Aerospace Engineering Sciences Department, University of Colorado, Boulder, CO 80309, USA

ARTICLE INFO

Article history:
Received 2 August 2013
Received in revised form 8 January 2014
Accepted 19 February 2014
Available online 26 February 2014

Keywords:
Compressive sampling
Sparse approximation
Polynomial chaos
Basis pursuit denoising (BPDN)
Weighted ℓ_1-minimization
Uncertainty quantification
Stochastic PDEs

ABSTRACT

This work proposes a method for sparse polynomial chaos (PC) approximation of high-dimensional stochastic functions based on non-adapted random sampling. We modify the standard ℓ_1-minimization algorithm, originally proposed in the context of compressive sampling, using *a priori* information about the decay of the PC coefficients, when available, and refer to the resulting algorithm as *weighted* ℓ_1*-minimization*. We provide conditions under which we may guarantee recovery using this weighted scheme. Numerical tests are used to compare the weighted and non-weighted methods for the recovery of solutions to two differential equations with high-dimensional random inputs: a boundary value problem with a random elliptic operator and a 2-D thermally driven cavity flow with random boundary condition.

1. Introduction

As we analyze engineering systems of increasing complexity, we must strategically confront the imperfect knowledge of the underlying physical models and their inputs, as well as the implied imperfect knowledge of a quantity of interest (QOI) predicted from these models. The understanding of outputs as a function of inputs in the presence of such uncertainty falls within the field of uncertainty quantification. The accurate quantification of the uncertainty of the QOI allows for the rigorous mitigation of both unfounded confidence and unnecessary diffidence in the anticipated QOI.

Probability is a natural mathematical framework for describing uncertainty, and so we assume that the system input is described by a vector of independent random variables, Ξ. If the random variable QOI, denoted by $u(\Xi)$, has finite variance, then the polynomial chaos (PC) expansion [1,2] is given in terms of the orthonormal polynomials $\{\psi_j(\Xi)\}$ as

$$u(\Xi) = \sum_{j=1}^{\infty} c_j \psi_j(\Xi). \tag{1}$$

A more detailed exposition on the use of PC expansion in this work is given in Section 2.2.

To identify the PC coefficients, c_j in (1), sampling methods including Monte Carlo simulation [3], pseudo-spectral stochastic collocation [4–7], or least-squares regression [8] may be applied. These methods for evaluating the PC coefficients are popular in that deterministic solvers for the QOI may be used without being adapted to the probability space. However, the

* Corresponding author.
E-mail address: alireza.doostan@colorado.edu (A. Doostan).

http://dx.doi.org/10.1016/j.jcp.2014.02.024
0021-9991/

standard Monte Carlo approach suffers from a slow convergence rate. Additionally, a major limitation to the use of the last two approaches above, in their standard form, is that the number of samples needed to approximate c_j generally increases rapidly with the dimension of the input uncertainty, i.e., the number of random variables needed to describe the input uncertainty. In particular, for asymptotically large dimensions, such growth is exponential, see, e.g., [9–13].

In this work, we use the Monte Carlo sampling method while considerably improving the accuracy of approximated PC coefficients (for the same number of samples) by exploiting the approximate sparsity of the coefficients c_j. As u has finite variance, the c_j in (1) necessarily converge to zero, and if this convergence is sufficiently rapid, then $u(\boldsymbol{\Xi})$ may be approximated by

$$\hat{u}(\boldsymbol{\Xi}) = \sum_{j\in\mathcal{C}} c_j \psi_j(\boldsymbol{\Xi}), \tag{2}$$

where the index set $\mathcal{C}$ has few elements. When this occurs we say that $\hat{u}$ is reconstructed from a sparse PC expansion, and that u admits an approximately sparse PC representation. By truncating the PC basis implied by (1) to P elements, we may perform calculations on the truncated PC basis. If we let $\boldsymbol{c}$ be a vector of c_j, for $j = 1, \dots, P$, then the approximate sparsity of the QOI (implied by the sparsity of $\boldsymbol{c}$) and the practical advantage of representing the QOI with a small number of basis functions motivate a search for an approximate $\boldsymbol{c}$ which has few non-zero entries [14–20]. We seek to achieve an accurate reconstruction with a small number of samples, and so look to techniques from the field of compressive sampling [21–29].

Let $\boldsymbol{\xi}$ represent a realization of $\boldsymbol{\Xi}$. We define $\boldsymbol{\Psi}$ as the matrix where each row corresponds to the row vector of P PC basis functions evaluated at sampled $\boldsymbol{\xi}$ with the corresponding $u(\boldsymbol{\xi})$ being an entry in the vector $\boldsymbol{u}$. We assume $N < P$ samples of $\boldsymbol{\xi}$, so that $\boldsymbol{\Psi}$ is $N \times P$, $\boldsymbol{c}$ is $P \times 1$, and $\boldsymbol{u}$ is $N \times 1$. Compressive sampling seeks a solution $\boldsymbol{c}$ with minimum number of non-zero entries by solving the optimization problem

$$\mathcal{P}_{0,\epsilon} \equiv \left\{ \arg\min_{\boldsymbol{c}} \|\boldsymbol{c}\|_0 \colon \|\boldsymbol{\Psi}\boldsymbol{c} - \boldsymbol{u}\|_2 \leqslant \epsilon \right\}. \tag{3}$$

Here $\|\boldsymbol{c}\|_0$ is defined as the number of non-zero entries of $\boldsymbol{c}$, and a solution to $\mathcal{P}_{0,\epsilon}$ directly provides an optimally sparse approximation in that a minimal number of non-zero entries are used to recover $\boldsymbol{u}$ to within ϵ in the ℓ_2 norm. In general, the cost of finding a solution to $\mathcal{P}_{0,\epsilon}$ grows exponentially in P [29]. To resolve this exponential dependence, the convex relaxation of $\mathcal{P}_{0,\epsilon}$ based on ℓ_1-minimization, also referred to as basis pursuit denoising (BPDN), has been proposed [21,22,24,23,29]. Specifically, BPDN seeks to identify $\boldsymbol{c}$ by solving

$$\mathcal{P}_{1,\epsilon} \equiv \left\{ \arg\min_{\boldsymbol{c}} \|\boldsymbol{c}\|_1 \colon \|\boldsymbol{\Psi}\boldsymbol{c} - \boldsymbol{u}\|_2 \leqslant \epsilon \right\} \tag{4}$$

using convex optimization algorithms [21,30–36]. In practice, $\mathcal{P}_{0,\epsilon}$ and $\mathcal{P}_{1,\epsilon}$ may have similar solutions, and the comparison of the two problems has received significant study, see, e.g., [29] and the references therein.

Note in (4) the constraint $\|\boldsymbol{\Psi}\boldsymbol{c} - \boldsymbol{u}\|_2 \leqslant \epsilon$ depends on the observed $\boldsymbol{\xi}$ and $u(\boldsymbol{\xi})$; not in general $\boldsymbol{\Xi}$ and $u(\boldsymbol{\Xi})$. As a result, $\boldsymbol{c}$ may be chosen to fit the input data, and not accurately approximate $u(\boldsymbol{\Xi})$ for previously unobserved realizations $\boldsymbol{\xi}$. To avoid this situation, we determine ϵ by cross-validation [16] as discussed in Section 2.5.

To assist in identifying a solution to (4), note that for certain classes of functions, theoretical analysis suggests estimates on the decay for the magnitude of the PC coefficients [37–39]. Alternatively, as we shall see in Section 4.2, such estimates may be derived by taking into account certain relations among physical variables in a problem. It is reasonable to use this *a priori* information to improve the accuracy of sparse approximations [40]. Moreover, even if this decay information is unavailable, each approximated set of PC coefficients may be considered as an initialization for the calculation of an improved approximation, suggesting an iterative scheme [41,40,42,43,18,20].

In this work, we explore the use of *a priori* knowledge of the PC coefficients as a weighting of ℓ_1 norm in BPDN in what is referred to as *weighted ℓ_1-minimization* (or *weighted BPDN*),

$$\mathcal{P}^{(\boldsymbol{W})}_{1,\epsilon} \equiv \left\{ \arg\min_{\boldsymbol{c}} \|\boldsymbol{W}\boldsymbol{c}\|_1 \colon \|\boldsymbol{\Psi}\boldsymbol{c} - \boldsymbol{u}\|_2 \leqslant \epsilon \right\}, \tag{5}$$

where $\boldsymbol{W}$ is a diagonal weight matrix to be specified. Previously, ℓ_1-minimization has been applied to solutions of stochastic partial differential equations with approximately sparse $\boldsymbol{c}$ [14,16,18,20], but these approximately sparse $\boldsymbol{c}$ include a number of small magnitude entries which inhibit the accurate recovery of larger magnitude entries. The primary goal of this work is to utilize *a priori* information about $\boldsymbol{c}$, in the form of estimates on the decay of its entries, to reduce this inhibition and enhance the recovery of a larger proportion of PC coefficients; in particular those of the largest magnitude. We provide theoretical results pertaining to the quality of the solution identified from the weighted ℓ_1-minimization problem $\mathcal{P}^{(\boldsymbol{W})}_{1,\epsilon}$.

1.1. Our contribution

This work specifically focuses on weights derived from *a priori* information for the anticipated solution vector as in Section 3.1. This *a priori* information, while not always available, may be produced from analytical convergence analysis as examined in Section 4.1 or scaling arguments as considered in Section 4.2. Additionally, as samples from a low-fidelity or