ARTICLE IN PRESS

Computers and Mathematics with Applications ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

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

Fast high-dimensional node generation with variable density

O. Vlasiuk [a,*], T. Michaels [a], N. Flyer [b], B. Fornberg [c]

[a] Center for Constructive Approximation, Department of Mathematics, Vanderbilt University, Nashville, TN 37240, USA
[b] Analytics and Integrated Machine Learning, National Center for Atmospheric Research, Boulder, CO 80305, USA
[c] Department of Applied Mathematics, University of Colorado, Boulder, CO 80309, USA

ARTICLE INFO

Article history:
Received 27 October 2017
Received in revised form 29 May 2018
Accepted 13 July 2018
Available online xxxx

Keywords:
Quasi-uniformity
Irrational lattice
Quasi-Monte Carlo methods
Riesz energy
Separation distance
Nearest neighbors

ABSTRACT

We present an algorithm for producing discrete distributions with a prescribed nearest-neighbor distance function. Our approach is a combination of quasi-Monte Carlo (Q-MC) methods and weighted Riesz energy minimization: the initial distribution is a stratified Q-MC sequence with some modifications; a suitable energy functional on the configuration space is then minimized to ensure local regularity. The resulting node sets are good candidates for building meshless solvers and interpolants, as well as for other purposes where a point cloud with a controlled separation-covering ratio is required. Applications of a three-dimensional implementation of the algorithm, in particular to atmospheric modeling, are also given.

1. Introduction

1.1. RBFs and meshless methods

In a number of important applications, usefulness of meshless methods in general, and of radial basis functions (RBFs) in particular, is well-known. They have found their way into high-dimensional interpolation, machine learning, spectral methods, vector-valued approximation and interpolation, just to name a few [1–5]. RBFs have multiple advantages, most importantly extreme flexibility in forming stencils (in the case of RBF-FD) and high local adaptivity; allowing spectral accuracy on irregular domains; the fact that the corresponding interpolation matrix (denoted by $\boldsymbol{A}$ below) is positive definite for several types of radial functions and does not suffer from instability phenomena characteristic of some of the alternative interpolation methods.

Applying RBF-FD stencils to building solvers requires an efficient way of distributing the *nodes* of basis elements in the domain, which can be either a solid or a surface. The tasks of modeling and simulation often call for massive numbers of nodes, so it is important to ensure that the distribution process is easily scalable. One further has to be able to place the RBFs according to a certain density, as a method of local refinement, for example, at the boundary, or in regions of special interest. Yet another challenge arises when it is necessary to deal with complex or non-smooth domains and/or surfaces.

Recall [6] that an RBF is a linear combination of the form

$$\mathcal{S}(\boldsymbol{x}) = \sum_{k=1}^{K} c_k \phi(\|\boldsymbol{x} - \boldsymbol{x}_k\|), \tag{1}$$

* Corresponding author.
E-mail addresses: oleksandr.vlasiuk@vanderbilt.edu (O. Vlasiuk), timothy.j.michaels@vanderbilt.edu (T. Michaels), flyer@ucar.edu (N. Flyer), bengt.fornberg@colorado.edu (B. Fornberg).

https://doi.org/10.1016/j.camwa.2018.07.026
0898-1221/

Please cite this article in press as: O. Vlasiuk, et al., Fast high-dimensional node generation with variable density, Computers and Mathematics with Applications (2018), https://doi.org/10.1016/j.camwa.2018.07.026.

where $\phi(\cdot)$ is a radial function, and $\boldsymbol{x}_k,\ k = 1, \ldots, K$, is a collection of pairwise distinct points in $\mathbb{R}^d$. A common choice of ϕ is the Gaussian $\phi(r) = e^{-(\epsilon r)^2}$, although one may also use $1/(1 + (\epsilon r)^2)$, $r^{2p}\log(r),\ p \in \mathbb{N}$, etc. In this discussion, we are not concerned with the distinctions between the different radial kernels, so the reader can assume that $\phi(r) = e^{-(\epsilon r)^2}$. In contrast to pseudospectral methods [3], RBF-FD approach means that to obtain a useful approximation of a function, or a differential operator, the nodes in expressions like (1) must be in the vicinity of the point $\boldsymbol{x}$, and therefore a large number of stencils are constructed throughout the underlying set. It is well-known that the matrix

$$\boldsymbol{A} = \begin{bmatrix} \phi(\|\boldsymbol{x}_1 - \boldsymbol{x}_1\|) & \phi(\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|) & \ldots & \phi(\|\boldsymbol{x}_1 - \boldsymbol{x}_K\|) \\ \phi(\|\boldsymbol{x}_2 - \boldsymbol{x}_1\|) & \phi(\|\boldsymbol{x}_2 - \boldsymbol{x}_2\|) & \ldots & \phi(\|\boldsymbol{x}_2 - \boldsymbol{x}_K\|) \\ \vdots & \vdots & & \vdots \\ \phi(\|\boldsymbol{x}_K - \boldsymbol{x}_1\|) & \phi(\|\boldsymbol{x}_K - \boldsymbol{x}_2\|) & \ldots & \phi(\|\boldsymbol{x}_K - \boldsymbol{x}_K\|) \end{bmatrix} \tag{2}$$

is positive definite if the nodes $\boldsymbol{x}_1 \ldots \boldsymbol{x}_K$ are all distinct [7], and so under this assumption there exist K-point RBF interpolants for any function data. A different question, however, is whether the matrix $\boldsymbol{A}$ will be well-conditioned: it is not the case, for example, when the nodes are placed on a lattice and $\epsilon \to 0$, [8]. The other extreme, having low regularity, also does not provide a reliable source of nodes, as can be seen on the example of the Halton sequence [8]. Furthermore, node clumping can lead to instability of PDE solvers, [3]. To avoid this, one must guarantee that the nodes are separated. In effect, generally the quasi-uniform node sets generated by the present algorithm, or, for example, the one constructed by the third and fourth authors [9], perform better than either lattice or the Halton sequence.

In many applications, one has to ensure that the distance from a node $\boldsymbol{x}$ to its nearest neighbor behaves approximately as a function of the position of the node [9]. Prescribing this function, $\rho(\boldsymbol{x})$, which we call the *radial density*, is a natural way to treat the cases when a local refinement is required in order to capture special features of the domain. In the present paper we will describe a method of node placement for which the actual distance to the nearest neighbor, denoted by $\Delta(\boldsymbol{x}) = \min_{\boldsymbol{x}' \neq \boldsymbol{x}} \|\boldsymbol{x}' - \boldsymbol{x}\|$, satisfies the above description. To summarize, we are interested in a procedure for obtaining discrete configurations inside a compact set that will:

- guarantee that $\Delta(\boldsymbol{x}) \asymp \rho(\boldsymbol{x})$ (that is, differ only up to a constant factor) for a given function $\rho(\boldsymbol{x})$ with a reasonably wide choice of ρ;
- be suitable for mesh-free PDE discretizations using RBFs, i.e., produce well-separated configurations without significant node alignment;
- result in quasi-uniform node distributions also on the surface boundaries of the domain;
- be computationally efficient, easily scalable, and suitable for parallelization.

1.2. Notation and layout

The bold typeface is reserved for vectors in $\mathbb{R}^d$. With few exceptions, letters of the Greek alphabet denote functions, calligraphic letters and Ω denote sets, and the regular Roman typeface is used for scalar variables. The symbolic notation employed throughout the paper is summarized in Table 1.

The paper is structured as follows: Section 2.1 outlines the RBF-FD method using Gaussian and Polyharmonic Spline kernels; Sections 2.2 and 2.3 introduce the two essential components of our approach, Riesz energy functionals and quasi-Monte Carlo methods. The main algorithm and its discussion are the subjects of Sections 3.1 and 3.2, respectively. Sections 4.1–4.3 offer applications of the algorithm; the corresponding run times are summarized in Section 4.4. Section 5.1 contains comparisons of the condition numbers of RBF-FD matrices with stencils on periodic Riesz minimizers, Halton nodes, and the Cartesian grid; Section 5.2 discusses the range of dimensions where the present method is applicable. The Appendix is dedicated to numerical experiments with the mean and minimal separation distance of Riesz minimizers and irrational lattices.

2. Choice of method

2.1. RBF-FD approximations

In this section we shall outline the common practices involving RBFs, in order to motivate the requirements that have to be imposed on the node distribution used in the respective computations. For a more in-depth discussion see one of [1,6,10,11]. A significant portion of the RBF approach hinges on the theory of positive definite functions.

Suppose we need to approximate a linear operator $\mathfrak{L}$ acting on sufficiently smooth functions supported on Ω, given locally by their values at the nodes $\boldsymbol{x}_k,\ k = 1, \ldots, K$. More specifically, we need to compute the value $\mathfrak{L}\psi(\boldsymbol{x}_0)$ for some fixed $\boldsymbol{x}_0 \in \Omega$ and a variable function ψ. A generalization of the standard [12] finite-difference (FD) approach consists in constructing weights $w_k,\ k = 1, \ldots, K$, that recover the value of $\mathfrak{L}$ at $\boldsymbol{x}_0$ in the form

$$\mathfrak{L}\mathcal{S}(\boldsymbol{x}_0) = \sum_{k=1}^{K} w_k \mathcal{S}(\boldsymbol{x}_k), \tag{3}$$

Please cite this article in press as: O. Vlasiuk, et al., Fast high-dimensional node generation with variable density, Computers and Mathematics with Applications (2018), https://doi.org/10.1016/j.camwa.2018.07.026.