Radial Basis Functions: Theory and Implementations

M. D. BUHMANN
University of Giessen

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published 2003

Printed in the United Kingdom at the University Press, Cambridge

Typeface Times 10/13 pt *System* LaTeX 2_ε [TB]

A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication data

Buhmann, M. D. (Martin Dietrich), 1963–
Radial basis functions : theory and implementations / Martin Buhmann.
p. cm. – (Cambridge monographs on applied and computational mathematics; 12)
Includes bibliographical references and index.
ISBN 0 521 63338 9
1. Radial basis functions. I. Title. II. Series.
QA223 .B84 2003
511′.42 – dc21 2002034983

ISBN 0 521 63338 9 hardback

Contents

1

Introduction

In the present age, when computers are applied almost anywhere in science, engineering and, indeed, all around us in day-to-day life, it becomes more and more important to implement mathematical functions for efficient evaluation in computer programs. It is usually necessary for this purpose to use all kinds of ‘approximations’ of functions rather than their exact mathematical form. There are various reasons why this is so. A simple one is that in many instances it is not possible to implement the functions exactly, because, for instance, they are only represented by an infinite expansion. Furthermore, the function we want to use may not be completely known to us, or may be too expensive or demanding of computer time and memory to compute in advance, which is another typical, important reason why approximations are required. This is true even in the face of ever increasing speed and computer memory availability, given that additional memory and speed will always increase the demand of the users and the size of the problems which are to be solved. Finally, the data that define the function may have to be computed interactively or by a step-by-step approach which again makes it suitable to compute approximations. With those we can then pursue further computations, for instance, or further evaluations that are required by the user, or display data or functions on a screen. Such cases are absolutely standard in mathematical methods for modelling and analysing functions; in this context, analysis can mean, e.g., looking for their stationary points with standard optimisation codes such as quasi-Newton methods.

As we can see, the applications of general purpose methods for functional approximations are manifold and important. One such class of methods will be introduced and is the subject area of this book, and we are particularly interested when the functions to be approximated (the approximands)

(a) depend on many variables or parameters,
(b) are defined by possibly very many data,

(c) and the data are ‘scattered’ in their domain.

The ‘radial basis function approach’ is especially well suited for those cases.

1.1 Radial basis functions

Radial basis function methods are the means to approximate the multivariate functions we wish to study in this book. That is, in concrete terms, given data in n dimensions that consist of data sites $\xi \in \mathbb{R}^n$ and ‘function values’ $f_\xi = f(\xi) \in \mathbb{R}$ (or $\mathbb{C}$ but we usually take $\mathbb{R}$), we seek an approximant $s\colon \mathbb{R}^n \to \mathbb{R}$ to the function $f\colon \mathbb{R}^n \to \mathbb{R}$ from which the data are assumed to stem. Here $n > 0$ is the dimension of underlying space and, incidentally, one often speaks – not quite correctly – of ‘data’ when referring just to the ξ. They can also be restricted to a domain $D \subset \mathbb{R}^n$ and if this D is prescribed, one seeks an approximation $s\colon D \to \mathbb{R}$ only. In the general context described in the introduction to this chapter, we consider $f(\xi)$ as the explicit function values we know of our f, which itself is unknown or at least unavailable for arbitrarily large numbers of evaluations. It could represent magnetic potentials over the earth’s surface or temperature measurements over an area or depth measurements over part of an ocean.

While the function f is usually not known in practice, for the purpose of (e.g. convergence) analysis, one has to postulate the existence of f, so that s and f can be compared and the quality of the approximation estimated. Moreover, some smoothness of f normally has to be required for the typical error estimates.

Now, given a linear space S of approximants, usually finite-dimensional, there are various ways to find approximants $s \in S$ to approximate the approximand (namely, the object of the approximation) f. In this book, the approximation will normally take place by way of interpolation, i.e. we explicitly require $s|_\Xi = f|_\Xi$, where $\Xi \subset \mathbb{R}^n$ is the discrete set of data sites we have mentioned above. Putting it another way, our goal is to interpolate the function between the data sites. It is desirable to be able to perform the interpolation – or indeed any approximation – without any further assumptions on the shape of Ξ, so that the data points can be ‘scattered’. But sometimes we assume $\Xi = (h\mathbb{Z})^n$, h a positive step size, $\mathbb{Z}$ the integers, for example, in order that the properties of the approximation method can more easily be analysed. We call this type of data distribution a square (cardinal) grid of step size h. This is only a technique for analysis and means no restriction for application of the methods to scattered ξ. Interpolants probably being the most frequent choice of approximant, other choices are nonetheless possible and used in practice, and they can indeed be very desirable such as least squares approximations or ‘quasi-interpolation’, a variant of interpolation, where s still depends in a simple way on f_ξ, $\xi \in \Xi$, while not necessarily matching each f_ξ exactly. We will come back to this type of approximation at many places in this book. We remark that if we know how to approximate a function $f\colon \mathbb{R}^n \to \mathbb{R}$ we can always approximate a vector-valued approximand, call it $F\colon \mathbb{R}^n \to \mathbb{R}^m$, $m > 1$, componentwise.

From these general considerations, we now come back to our specific concepts for the subject area of this monograph, namely, for radial basis function approximations the approximants s are usually finite linear combinations of translates of a radially symmetric basis function, say $\phi(\|\cdot\|)$, where $\|\cdot\|$ is the Euclidean norm. Radial symmetry means that the value of the function only depends on the Euclidean distance of the argument from the origin, and any rotations thereof make no difference to the function value.

The translates are along the points $\xi \in \Xi$, whence we consider linear combinations of $\phi(\|\cdot - \xi\|)$. So the data sites enter already at two places here, namely as the points where we wish to match interpolant s and approximand f, and as the vectors by which we translate our radial basis function. Those are called the centres, and we observe that their choice makes the space S dependent on the set Ξ. There are good reasons for formulating the approximants in this fashion used in this monograph.

Indeed, it is a well-known fact that interpolation to arbitrary data in more than one dimension can easily become a singular problem unless the linear space S from which s stems depends on the set of points Ξ – or the Ξ have only very restricted shapes. For any fixed, centre-independent space, there are some data point distributions that cause singularity.

In fact, polynomial interpolation is the standard example where this problem occurs and we will explain that in detail in Chapter 3. This is why radial basis functions always define a space $S \subset C(\mathbb{R}^n)$ which depends on Ξ. The simplest example is, for a finite set of centres Ξ in $\mathbb{R}^n$,

$$(1.1)\qquad S = \left\{ \sum_{\xi \in \Xi} \lambda_\xi \|\cdot - \xi\| \;\middle|\; \lambda_\xi \in \mathbb{R} \right\}.$$

Here the ‘radial basis function’ is simply $\phi(r) = r$, the radial symmetry stemming from the Euclidean norm $\|\cdot\|$, and we are shifting this norm in (1.1) by the centres ξ.

More generally, radial basis function spaces are spanned by translates

$$\phi(\|\cdot - \xi\|), \qquad \xi \in \Xi,$$

where $\phi\colon \mathbb{R}_+ \to \mathbb{R}$ are given, continuous functions, called radial basis functions. Therefore the approximants have the general form

$$s(x) = \sum_{\xi \in \Xi} \lambda_\xi \phi(\|\cdot - \xi\|), \qquad x \in \mathbb{R}^n,$$

with real coefficients λ_ξ.

Other examples that we will encounter very often from now on are $\phi(r) = r^2 \log r$ ('thin-plate splines'), $\phi(r) = \sqrt{r^2 + c^2}$ (c a positive parameter, 'multiquadrics'), $\phi(r) = e^{-\alpha r^2}$ (α a positive parameter, 'Gaussian'). As the later analysis will show, radial symmetry is not the most important property that makes these functions such suitable choices for approximating smooth functions as they are, but rather their smoothness and certain properties of their Fourier transform. Nonetheless we bow to convention and speak of radial basis functions even when we occasionally consider general n-variate $\phi\colon \mathbb{R}^n \to \mathbb{R}$ and their translates $\phi(\cdot - \xi)$ for the purpose of approximation. And, at any rate, most of these basis functions that we encounter in theory and practice *are* radial. This is because it helps in applications to consider genuinely radial ones, as the composition with the Euclidean norm makes the approach technically in many respects a univariate one; we will see more of this especially in Chapter 4. Moreover, we shall at all places make a clear distinction between considering general n-variate $\phi\colon \mathbb{R}^n \to \mathbb{R}$ and radially symmetric $\phi(\|\cdot\|)$ and carefully state whether we use one or the other in the following chapters.

Unlike high degree spline approximation with scattered data in more than one dimension, and unlike the polynomial interpolation already mentioned, the interpolation problem from the space (1.1) is *always* uniquely solvable for sets of distinct data sites ξ, and this is also so for multiquadrics and Gaussians. For multivariate polynomial spline spaces on nongridded data it is up to now not even possible in general to find the exact dimension of the spline space! Thus we may very well be unable to interpolate uniquely from that spline space. Only several upper and lower bounds on the spatial dimension are available. There exist radial basis functions ϕ of compact support, where there *are* some restrictions so that the interpolation problem is nonsingular, but they are only simple bounds on the dimension n of $\mathbb{R}^n$ from where the data sites come. We will discuss those radial basis functions of compact support in Chapter 6 of this book.

Further remarkable properties of radial basis functions that render them highly efficient in practice are their easily adjustable smoothness and their powerful convergence properties. To demonstrate both, consider the ubiquitous multiquadric function which is infinitely often continuously differentiable for $c > 0$ and only continuous for $c = 0$, since in the latter case $\phi(r) = r$ and

$\phi(\|\cdot\|)$ is the Euclidean norm as considered in (1.1) which has a derivative discontinuity at zero. Other useful radial basis functions of any given smoothness are readily available, even of compact support, as we have just mentioned. Moreover, as will be seen in Chapter 4, on $\Xi = (h\mathbb{Z})^n$, e.g. an approximation rate of $O(h^{n+1})$ is obtained with multiquadrics to suitably smooth f. This is particularly remarkable because the convergence rate increases linearly with dimension, and, at any rate, it is very fast convergence indeed. Of course, the amount of work needed (e.g. the number of centres involved) for performing the approximation also increases at the same rate. Sometimes, even exponential convergence orders are possible with multiquadric interpolation and related radial basis functions.

1.2 Applications

Consequently, it is no longer a surprise that in many applications, radial basis functions have been shown to be most useful. Purposes and applications of such approximations and in particular of interpolation are manifold. As we have already remarked, there are many applications especially in the sciences and in mathematics. They include, for example, mappings of two- or three-dimensional images such as portraits or underwater sonar scans into other images for comparison. In this important application, interpolation comes into play because some special features of an image may have to be preserved while others need not be mapped exactly, thus enabling a comparison of some features that may differ while at the same time retaining others. Such so-called ‘markers’ can be, for example, certain points of the skeleton in an X-ray which has to be compared with another one of the same person, taken at another time. The same structure appears if we wish to compare sonar scans of a harbour at different times, the rocks being suitable as markers this time. Thin-plate splines turned out to be excellent for such very practical applications (Barrodale and Zala, 1999).

Measurements of potential or temperature on the earth’s surface at ‘scattered’ meteorological stations or measurements on other multidimensional objects may give rise to interpolation problems that require the aforementioned scattered data. Multiquadric approximations are performing well for this type of use (Hardy, 1990).

Further, the so-called track data are data sites which are very close together on nearly parallel lines, such as can occur, e.g., in measurements of sea temperature with a boat that runs along lines parallel to the coast. So the step size of the measurements is very small along the lines, but the lines may have a distance of 100 times that step size or more. Many interpolation algorithms fail on such awkward distributions of data points, not so radial basis function (here multiquadric) methods (Carlson and Foley, 1992).

The approximation to so-called learning situations by neural networks usually leads to very high-dimensional interpolation problems with scattered data. Girosi (1992) mentions radial basis functions as a very suitable approach to this, partly because of their availability in arbitrary dimensions and of their smoothness.

A typical application is in fire detectors. An advanced type of fire detector has to look at several measured parameters such as colour, spectrum, intensity, movement of an observed object from which it must decide whether it is looking at a fire in the room or not. There is a learning procedure before the implementation of the device, where several prescribed situations (these are the data) are tested and the values zero (no fire) and one (fire) are interpolated, so that the device can 'learn' to interpolate between these standard situations for general situations later when it is used in real life.

In another learning application, the data come from the raster of a screen which shows the reading of a camera that serves as the eye of a robot. In this application, it is immediately clear why we have a high-dimensional problem, because each point on the square raster represents one parameter, which gives a million points even on a relatively low resolution of 1000 by 1000. The data come from showing objects to the robot which it should recognise as, for instance, a wall it should not run into, or a robot friend, or its human master or whatever. Each of these situations should be interpolated and from the interpolant the robot should then be able to recognise other, similar situations as well. Invariances such as those objects which should be recognised independently of angle etc. are also important in measurements of neural activity in the brain, where researchers aim to recognise those activities of the nerve cells that appear when someone is looking at an object and which are invariant of the angle under which the object is looked at. This is currently an important research area in neuro-physics (Eckhorn, 1999, Kremper, Schanze and Eckhorn 2002) where radial basis functions appear often in the associated physics literature. See the above paper by Eckhorn for a partial list.

The numerical solution of partial differential equations also enters into the long list of mathematical applications of radial basis function approximation. In the event, Pollandt (1997) used them to perform approximations needed in a multidimensional boundary element method to solve nonlinear elliptic PDEs on a domain Ω, such as $\Delta u_\ell = p_\ell(u(x), x)$, $x \in \Omega \subset \mathbb{R}^n$, $\ell = 1, \ldots, N$, with Dirichlet boundary conditions $u_\ell|_{\partial\Omega} = q_\ell$, when $u = (u_1, \ldots, u_N)^T$ are suitably smooth functions and p_ℓ are multivariate polynomials. Here, Δ denotes the Laplace operator. The advantage of radial basis functions in this application is the already mentioned convergence power, in tandem with easy formulation of interpolants and quasi-interpolants and their introduction into the PDE. Moreover, especially for boundary element methods, it is relevant that several radial basis functions are Green's functions of elliptic differential operators, i.e. the elliptic differential operator applied to them including the composition with the ubiquitous Euclidean norm yields the Dirac δ-operator.

The same reasons led Sonar (1996) to use radial basis functions for the local reconstruction of solutions within algorithms which solve numerically hyperbolic conservation laws. It was usual to employ low order polynomial approximation (mostly linear) for this purpose so far, but it turned out that radial basis functions, especially thin-plate splines, help to improve the accuracy of the finite volume methods notably to solve the hyperbolic equations, because of their ability to approximate locally ('recover' in the language of hyperbolic conservation laws) highly accurately.

They appear to be remarkably resilient against irregular data distributions, for not only track data but also those that occur, for instance, when local models are made for functions whose stationary points (or extrema) are sought (Powell, 1987). This is problematic because algorithms that seek such points will naturally accumulate data points densely near the stationary point, where now an approximation is made, based on those accumulated points, to continue with the approximant instead of the original function (which is expensive to evaluate). Furthermore, it turned out to be especially advantageous for their use that radial basis functions have a variation-diminishing property which is explained in Chapter 5. Thin-plate splines provide the most easily understood variant of that property and thus they were used for the first successful experiments with radial basis functions for optimisation algorithms. Not only do the variation-diminishing properties guarantee a certain smoothness of the approximants, but they are tremendously helpful for the analysis because many concepts of orthogonal projections and norm-minimising approximants can be used in the analysis. We shall do so often in this book.

In summary, our methods are known from practice to be good and general purpose approximation and interpolation techniques that can be used in many instances, where other methods are unlikely to deliver useful results or fail completely, due to singular interpolation matrices or too high dimensionality. The methods are being applied widely, and important theoretical results have been found that support the experimental and practical observations, many of which will enter into this book. Among them are the exceptional accuracy that can be obtained, when interpolating smooth functions.

Thus the purpose of this book is to demonstrate how well radial basis function techniques work and why, and to summarise and explain efficient implementations and applications. Moreover, the analysis presented in this work will allow a user to choose which radial basis functions to use for his application based on its individual merits. This is important because the theory itself, while mathematically stimulating, stands alone if it does not yield itself to support practical use.

1.3 Contents of the book

We now outline the contents and intents of the following chapters. The next chapter gives a brief summary of the schemes including precise mathematical details of some specific methods and aspects, so that the reader can get sufficient insight into the radial basis function approach and its mathematical underpinning to understand (i) the specifics of the methods and (ii) the kind of mathematical analysis typically needed. This may also be the point to decide whether this approach is suitable for his needs and interests and whether he or she wants to read on to get the full details in the rest of the book or, e.g., just wants to go directly to the chapter about implementations.

Chapter 3 puts radial basis functions in the necessary, more general context of multivariate interpolation and approximation methods, so that the reader can compare and see the ‘environment’ of the book. Especially splines, Shepard’s method, and several other widely used (mostly interpolation) approaches will be reviewed briefly in that chapter. There is, however, little on practice and implementations in Chapter 3. It is really only a short summary and not comprehensive.

Chapter 4 introduces the reader to the very important special case of $\Xi = (h\mathbb{Z})^n$, $h > 0$, i.e. radial basis functions on regularly spaced (integer) grids. This was one of the first cases when their properties were explicitly and comprehensively analysed and documented, because the absence of boundaries and the periodicity of Ξ allow the application of powerful analysis tools such as Fourier transforms, Poisson summation formula etc. While the analysis is easier, it still gives much insight into the properties of the functions and the spaces generated by them, such as unique existence of interpolants, conditioning of interpolation matrices and exactness of approximations to polynomials of certain degrees, and, finally but most importantly, convergence theorems. Especially the latter will be highly relevant to later chapters of the book. Moreover, several of the results on gridded Ξ will be seen to carry through to scattered data, so that indeed the properties of the spaces generated by translates of radial basis functions were documented correctly. Many of the results are shown to be best possible, too, that is, they explicitly give the best possible convergence results.

The following chapter, 5, generalises the results of Chapter 4 to scattered data, confirming that surprisingly many *approximation order results* on gridded data carry through with almost no change, whereas, naturally, new and involved proofs are needed. One of the main differences is that there are usually finitely many data for this setting and, of course, there are boundaries of the domain wherein the $\xi \in \Xi$ reside, which have to be considered. It is usual that there are less striking convergence results in the presence of boundaries and this is what we will find there as well.

This Chapter 5, dealing with the many and deep theorems that have been established concerning the convergence rates of approximations, is the core of the book. This is because, aside from the existence and uniqueness theorems about interpolation, convergence of the methods is of utmost importance in applications. After all, the various rates of convergence that can be achieved are essential to the choice of a method and the interpretation of its results. Besides algebraic rates of convergence that are related to the polynomial exactness results already mentioned, the aforementioned spectral rates are discussed.

In Chapter 6, radial basis functions with compact support are constructed. They are useful especially when the number of data or evaluations of the interpolant is massive so that any basis functions of global support incur prohibitive costs for evaluation. Many of those radial basis functions are piecewise polynomials, and all of them have similar nonsingularity properties for the interpolation problem to the ones we have mentioned before. Moreover, radial basis functions with compact support are suitable for, and are now actually used in, solving linear partial differential equations by Galerkin methods. There, they provide a suitable replacement for the standard piecewise polynomial finite elements. It turns out that they can be just as good as means for approximation, while not requiring any triangulation or mesh, so they allow meshless approximation which is easier when the amount of data has to be continuously enlarged or made smaller. That is often the case when partial differential equations are solved numerically. By contrast, finite elements can be difficult to compute in three or more dimensions for scattered data due to the necessity of triangulating the domain before using finite elements and due to the complicated spaces of piecewise polynomials in more than two dimensions.

While many such powerful theoretical results exist for radial basis functions, the implementation of the methods is nontrivial and requires careful attention. Thus Chapter 7 describes several modern techniques that have been developed to implement the approach, evaluate and compute interpolants fast and efficiently, so that real-time rendering of surfaces that interpolate the data is possible now, for example. The methods we describe are iterative and they include so-called particle methods, efficient preconditioners and the Beatson–Faul–Goodsell–Powell (BFGP) algorithm of local Lagrange bases.

As outlined above, the principal application of radial basis functions is clearly with interpolation. This notwithstanding, least squares methods are frequently asked for, especially because often in applications, data are inaccurate, too many, and/or need smoothing. Hence Chapter 8 is devoted to least squares approaches both using the standard Euclidean least squares setting and with the so-called Sobolev inner products. Existence and convergence questions will be considered as well as, briefly, implementation.

Closely related to the least squares problem, whose solution is facilitated by computing orthogonal or orthonormal bases of the radial basis function spaces in advance, are 'wavelet expansions' by radial basis functions. In these important wavelet expansions, the goal is to decompose a given function simultaneously into its local parts in space *and* in frequency. The purpose of this can be the analysis, approximation, reconstruction, compression, or filtering of functions and signals. In comparison with the well-known Fourier analysis we can, e.g., tell from a wavelet expansion *when* a frequency appears in a melody, say, and not just that it appears and with what amplitude. This is what we call *localness*, not only in frequency for the wavelet expansions. Radial basis functions are bound to be useful for this because of their approximational efficacy. After all, the better we can approximate from a space, the fewer coefficients are needed in expansions of functions using bases of that space. All this is detailed, together with several examples of radial basis (especially multiquadric) wavelets, in Chapter 9.

Chapter 10 concerns the most recent and topical results in review form and an outlook and ideas towards further, future research. Many aspects of these tools are studied in research articles right now and in Chapter 10 we attempt to catch up with the newest work. Of course this can only be discussed very briefly. Several important questions are still wide open and we will outline some of those.

We conclude with an extensive bibliography, our principal aim being to provide a good account of the state of the art in radial basis function research. Of course not all aspects of current or past interest in radial basis functions can be covered within the scope of a book of this size but the aim is at least to provide up to date references to those areas that are not covered. We also give a commentary on the bibliography to point the reader to other interesting results that are not otherwise in this book on one hand, and to comment on generalisations, other points of view etc. on those results that are.

Now, in the following chapter, we give the already mentioned summary of some aspects of radial basis functions in detail in order to exemplify the others.