Available online at www.sciencedirect.com
ScienceDirect
Applied Mathematical Modelling 33 (2009) 1573–1583
APPLIED MATHEMATICAL MODELLING
www.elsevier.com/locate/apm

On the selection of the most adequate radial basis function

Humberto Rocha *

Portuguese Catholic University, Pólo de Viseu, 3504-505 Viseu, Portugal

Received 13 September 2007; received in revised form 7 January 2008; accepted 13 February 2008
Available online 10 March 2008

Abstract

Radial basis function (RBF) methods can provide excellent interpolants for a large number of poorly distributed data points. For any finite data set in any Euclidean space, one can construct an interpolation of the data by using RBFs. However, RBF interpolant trends between and beyond the data points depend on the RBF used and may exhibit undesirable trends using some RBFs while the trends may be desirable using other RBFs. The fact that a certain RBF is commonly used for the class of problems at hand, previous good behavior in that (or other) class of problems, and bibliography, are just some of the many valid reasons given to justify a priori selection of RBF. Even assuming that the justified choice of the RBF is most likely the correct choice, one should nonetheless confirm numerically that, in fact, the most adequate RBF for the problem at hand is the RBF chosen a priori. The main goal of this paper is to alert the analyst as to the danger of a priori selection of RBF and to present a strategy to numerically choose the most adequate RBF that better captures the trends of the given data set. The wing weight data fitting problem is used to illustrate the benefits of an adequate choice of RBF for each given data set.

Keywords: Radial basis function; Data fitting; Cross-validation; Principal component regression

1. Introduction

Radial basis function (RBF) methods can provide excellent interpolants for high dimensional data sets of poorly distributed data points (scarce and unevenly distributed points). For any finite data set in any Euclidean space, one can construct an interpolation of the data by using RBFs, even if the data points are unevenly and sporadically distributed in a high dimensional Euclidean space. There is a wide range of applications where RBF interpolation methods can be successfully applied, such as aeronautics, meteorology and medical imaging (see [1–5]). However, RBF interpolant trends between and beyond the data points depend on the RBF used and may exhibit undesirable trends using some RBFs while the trends may be desirable using other RBFs.

The fact that a certain RBF is commonly used for the class of problems at hand, previous good behavior in that (or other) class of problems, and bibliography, are just some of the many valid reasons given to justify a

* Tel.: +351 239109410.
E-mail address: hrocha@mat.uc.pt

0307-904X/$ - see front matter
doi:10.1016/j.apm.2008.02.008

priori selection of RBF. Even assuming that the justified choice of the RBF is most likely the correct choice, i.e., that the RBF picked is the most adequate and the one that produces the "best" model, one should nonetheless confirm numerically that, in fact, the most adequate RBF for the problem at hand is the RBF chosen a priori.

The main goal of this paper is to alert the analyst as to the danger of a priori choice of RBF and to present a strategy to numerically choose the most adequate RBF that better captures the trends of the given data set. The wing weight data fitting problem is used to illustrate the benefits of an adequate choice of RBF for each given data set.

The paper is organized as follows. Section 2 gives a brief description of RBF interpolation problems. Section 3 introduces cross-validation for model parameter tuning. In Section 4, wing weight data fitting is used to illustrate the benefits of an adequate choice of RBF for each given data set. Section 5 includes the concluding remarks.

2. RBF interpolation problems

Let $f(\mathbf{x})$ be a real-valued function of the input vector $\mathbf{x}$ defined on a subset Ω of $\mathbb{R}^n$ such that the value of f at N input vectors $\mathbf{x}^j$, $j = 1, \ldots, N$, $f(\mathbf{x}^j)$, is given. The goal is to construct an estimation model $g(\mathbf{x})$ such that $g(\mathbf{x}^j) = f(\mathbf{x}^j)$ for $j = 1, \ldots, N$. The interpolation requirement can be satisfied by RBF interpolation.

Interpolation functions generated from a RBF $\varphi(t)$ can be represented in the following form:

$$g(\mathbf{x}) = \sum_{j=1}^{N} \alpha_j \varphi(\|\mathbf{x} - \mathbf{x}^j\|), \qquad (1)$$

where $||\mathbf{x} - \mathbf{x}^j||$ denotes the parameterized distance between $\mathbf{x}$ and $\mathbf{x}^j$ defined as

$$||\mathbf{x} - \mathbf{x}^j|| = \sqrt{\sum_{i=1}^{n} |\theta_i|(x_i - x_i^j)^2},$$

and $\theta_1, \ldots, \theta_n$ are scalars (see [1]).

The most popular examples of RBF [6–8] are cubic spline $\varphi(t) = t^3$, thin plate spline $\varphi(t) = t^2 \ln t$, multiquadric $\varphi(t) = \sqrt{1 + t^2}$, and Gaussian $\varphi(t) = \exp(-t^2)$ (see Fig. 1). These RBFs can be used to model cubic, almost quadratic, and linear growth rates, as well as exponential decay, of the response for trend predictions.

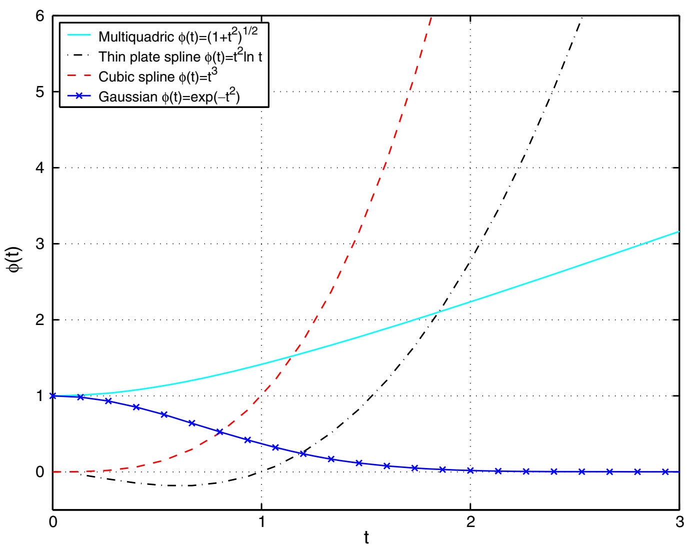

Fig. 1. Graphs of commonly used radial basis functions.