LOCAL MODIFICATIONS TO RADIAL BASIS NETWORKS

I. J. Anderson

School of Computing and Mathematics, University of Huddersfield, Queensgate, Huddersfield HD1 3DH, UK.

The form of a radial basis network is a linear combination of translates of a given radial basis function, \( \phi(r) \). The radial basis method involves determining the values of the unknown parameters within the network given a set of inputs, \( \{x_k\} \), and their corresponding outputs, \( \{f_k\} \). It is usual for some of the parameters of the network to be fixed. If the positions of the centres of the basis functions are known and constant, the radial basis problem reduces to a standard linear system of equations and many techniques are available for calculating the values of the unknown coefficients efficiently. However, if both the positions of the centres and the values of the coefficients are allowed to vary, the problem becomes considerably more difficult. A highly non-linear problem is produced and solved in an iterative manner. An initial guess for the best positions of the centres is made and the coefficients for this particular choice of centres are calculated as before. For each iteration, a small change to the position of the centres is made in order to improve the quality of the network and the values of the coefficients for these new centre positions are determined. The overall algorithm is computationally expensive and here we consider ways of improving the efficiency of the method by exploiting the local stability of the thin plate spline basis function. At each step of the iteration, only a small change is made to the positions of the centres and so we can reasonably expect that there is only a small change to the values of the corresponding coefficients. These small changes are estimated using local modifications.

1 Introduction

We consider the thin plate spline basis function

\[ \phi(r) = r^2 \log r. \]

This basis function has not been as popular as the Gaussian, \( \phi(r) = \exp(-r^2/2\sigma) \), mainly due to its unbounded nature. The thinking behind this is that it is desirable to use a basis function that has near compact support so that the approximating function behaves in a local manner.

A suitable definition of local behaviour for an approximating function of the form given in equation (1) is as follows. The \( i \)th coefficient, \( c_i \), of the approximant should be related to the values \( \{f_k\} \) for values of \( k \) where \( \|x_k - \lambda_i\| \) is small. Thus, the value of the coefficient should be influenced only by the data ordinates whose abscissae values are close to the centre of the corresponding basis function. From this definition one deduces that it is advisable to use basis functions that decay, hence the desire for basis functions with near compact support. However many authors have shown that this deduction is unfounded and it is in fact easier to produce the properties of local behaviour by using unbounded functions [4, 2, 3, 1].

1.1 The Approximation Problem

Given a set of \( m \) data points, \( (x_k, f_k) \), for \( k = 1, 2, \ldots, m \), it is possible to produce an approximation of the form

\[ f(x) = \sum_{i=1}^{n} c_i \phi(||x - \lambda_i||), \quad (1) \]
such that
\[ f(x_k) \approx f_k. \]

It is usual for this approximation problem to be solved in the least-squares sense. There are two important categories for such an approximation. The easiest approach is to consider the positions of the centres to be fixed, in which case the approximation problem is linear and is therefore reasonably efficient to solve. The alternative is to allow the positions of the centres to vary or indeed for the number of centres to change. For example, once a preliminary approximation has been completed it is worth examining the results to determine the quality of the approximation. It may be decided that one (or more) regions are unsatisfactory and so extra basis functions or an adjustment of the centres of the current basis functions would be suitable.

Under such circumstances it is usual to recalculate the new coefficients for the approximation problem with respect to the new positions of the centres. Repeating this process too often can be computationally expensive and it is more appropriate to modify the current approximation to take into account the small changes that have been made to the centres.

2 Local Stability of the Thin Plate Spline

This local behaviour of the thin plate spline is demonstrated by the use of an example that consists of approximating \( m = 6,864 \) data points using \( n = 320 \) basis functions. The centres for these basis functions are produced using a clustering algorithm and the data are fitted in the least-squares sense. Formally, let \( I \) be the set of indices for the basis functions, \( \{1, 2, \ldots, n\} \). We wish to calculate the values of the coefficients \( \{c_i\} \) that minimize the \( \ell_2 \)-norm of the residual vector \( e \) which has the components

\[ e_k = f_k - \sum_{i \in I} c_i \phi(||x - \lambda_i||), \]

for \( k = 1, 2, \ldots, m \). The data, centres and the resulting approximant are shown in Figure 1.