Applied Mathematics and Computation 247 (2014) 1129–1138

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

On the approximation of strongly convex functions by an upper or lower operator

Osama Yousif Mohammed Alabdali, Allal Guessab *

Laboratoire de Mathématiques et de leurs Applications, UMR CNRS 4152, Université de Pau et des Pays de l'Adour, 64000 Pau, France

ARTICLE INFO

Keywords:
Approximation
Convexity
Error estimates
Lipschitz gradients
Strongly convex functions

ABSTRACT

The aim of this paper is to find a convenient and practical method to approximate a given real-valued function of multiple variables by linear operators, which approximate all *strongly* convex functions from above (or from below). Our main contribution is to use this additional knowledge to derive sharp error estimates for continuously differentiable functions with Lipschitz continuous gradients. More precisely, we show that the error estimates based on such operators are always controlled by the Lipschitz constants of the gradients, the convexity parameter of the strong convexity and the error associated with using the quadratic function, see Theorems 3.1 and 3.3. Moreover, assuming the function, we want to approximate, is also strongly convex, we establish sharp upper as well as lower refined bounds for the error estimates, see Corollaries 3.2 and 3.4. As an application, we define and study a class of linear operators on an arbitrary polytope, which approximate strongly convex functions from above. Finally, we present a numerical example illustrating the proposed method.

1. Some background and motivation

Let $\Omega \subset \mathbb{R}^d$ be a nonempty compact convex set and let $\phi : \Omega \to \mathbb{R}$ be a given function. We would like to find an easier and good approximation to compute ϕ. We sometimes know beforehand that the function ϕ satisfies various known structural and regularity properties. For example, it may be known that ϕ has some additional kind of convexity, therefore we would wish to use this information in order to get a good approximation of ϕ. Approximating an arbitrary function is, in general, very difficult, but if we restrict our attention to the class of strongly convex functions and if the linear operator, we wish to use, approximates all *strongly* convex functions from above (or from below) then things become simpler. The strongly convex functions are used widely in economic theory (see [15]), and are also central to optimization theory (see [11]). Indeed, in the framework of function minimization, this convexity notion has important and well-known implications. As we will see, the key advantage of dealing with such an operator is that an estimate of its approximation error is always controlled by the error associated with using the quadratic function.

In order to illustrate this idea more precise, we start by describing briefly a specific one-dimensional example, since its simplicity allows us to analyze all the necessary steps through very simple computation. Suppose that σ is a fixed nonnegative real number. In the univariate approximation, say on an interval $[a, b]$, a simple way of approximating a given real σ-strongly function $\phi : [a, b] \to \mathbb{R}$ is first to choose a partition $P := \{x_0, x_1, \ldots, x_n\}$ of the interval $[a, b]$, such that $a = x_0 < x_1 < \cdots < x_n = b$, and then to fit to ϕ using a linear interpolant B_n at these points in such a way that:

* Corresponding author.
E-mail addresses: alsltan2006@yahoo.com (O.Y. Mohammed Alabdali), allal.guessab@univ-pau.fr (A. Guessab).

http://dx.doi.org/10.1016/j.amc.2014.09.007
0096-3003/

1. The domain of B_n is the interval $[a, b]$;
2. B_n is a linear polynomial on each subinterval $[x_i, x_{i+1}]$;
3. B_n is continuous on $[a, b]$ and it interpolates the data, that is, $B_n(x_i) = \phi(x_i),\ i = 0, \ldots, n$.

This is a convenient class of interpolants because every such interpolant can often be written for all $i = 0, \ldots, n-1$ in a barycentric form:

$$B_n[\phi](x) = \frac{x_{i+1} - x}{x_{i+1} - x_i}\phi(x_i) + \frac{x - x_i}{x_{i+1} - x_i}\phi(x_{i+1}), \quad (x \in [x_i, x_{i+1}]). \tag{1}$$

One of the main features of the usual linear interpolant, in its simplest form (1), is that the error in approximating the quadratic function $(.)^2$ by B_n is simply given by:

$$B_n[(.)^2](x) - x^2 = (x - x_i)(x_{i+1} - x), \quad (x \in [x_i, x_{i+1}]), \quad (x \in [x_i, x_{i+1}])$$

and also that B_n approximates all σ-strongly convex functions from above. More precisely, B_n satisfies for any σ-strongly convex function the following estimates:

$$\frac{\sigma}{2}(x - x_i)(x_{i+1} - x) \leqslant B_n[f](x) - f(x), \quad (x \in [x_i, x_{i+1}]).$$

Moreover, as it can be derived from our multivariate general results, see Remark 4.5, if we also know that the first derivative of f is a Lipschitz function with a local Lipschitz constant $L_i(f')$ in the subintervals $[x_i, x_{i+1}]$, then the error $B_n[f] - f$ can often be estimated at any $x \in [x_i, x_{i+1}]$ as:

$$\frac{\sigma}{2}(x - x_i)(x_{i+1} - x) \leqslant B_n[f](x) - f(x) \leqslant \frac{L_i(f')}{2}(x - x_i)(x_{i+1} - x). \tag{2}$$

Hence, the lower and upper bounds of the approximation error for this class of functions can be controlled by the Lipschitz constants of the first derivatives, the convexity parameter (of the strong convexity) and the error associated with using the quadratic function. It should be noted that equalities in (2) are attained for all σ-strongly convex functions of the form

$$f(x) = a(x) + \frac{\sigma}{2}x^2, \tag{3}$$

where $a(\cdot)$ is any affine function. Therefore, in this sense, the error estimates (2) are sharp for the class of σ-strongly convex functions having Lipschitz continuous first derivatives. This provides the starting point of the forthcoming results.

This paper deals with the problem of approximation of functions of multiple variables by using linear operators, which approximate from above (or from below) all strongly convex functions with Lipschitz-continuous gradients. Geometrically, if a function f belongs to such a class, then its gradient ∇f cannot change too quickly and it cannot change too slowly either. Functions satisfying these conditions are widely used in the optimization literature, we refer to Nesterov's book [11]. A natural question is: can these functions be well approximated by simpler functions and how?

There are several papers investigating various methods to approximate arbitrary functions, very little research has been done subject to some kind of additional convexity assumption. For instance, if some smoothness is allowed for the function, which is to be approximated, say $C^2(\Omega)$, this will play a crucial role in the determination of the "best" (or "optimal") cubature formulas, see [1–10]. This article builds on the previous work [3,4], where a theoretical framework for approximating $C^2(\Omega)$-convex functions was developed.

The motivation for such an approach is that the general sharp error estimates, that we derive, permit us to study a multivariate version defined on an arbitrary (convex) polytope of the univariate interpolation operators given by (1). Throughout the paper, a linear operator is said to be upper (resp. lower) operator for strongly convex functions, if it approximates from above (resp. from below) strongly convex functions.

The paper is organized as follows: In Section 2 we state some definitions and some of the basic properties and facts about strongly convex functions. The main theorems of Section 3 establish, in terms of sharp error estimates, simple and elegant characterizations of upper or lower approximation operators for strongly convex functions with Lipschitz-continuous gradients. In this way, we offer sharp error estimates which only depend on the Lipschitz constants of the gradients, the convexity parameter (of the strong convexity), and the error associated with using the quadratic function, see Theorems 3.1 and 3.3. This also allows us to recover and extend the simple approach of [3], which presented in the case where the functions are only convex. A particularly interesting situation arises, when the function, we want to approximate, is also strongly convex. In this special case, we establish sharp upper as well as lower refined bounds for the error estimates, see Corollaries 3.2 and 3.4. In Section 4, we will introduce and study a multivariate version defined on an arbitrary polytope of the univariate interpolation operators given by (1). Finally, Section 5 will provide a numerical example to illustrate the efficiency of this approach.

2. Some inequalities involving strongly convex functions

In this section, we state some definitions and properties of strongly convex functions, which are very useful in the proofs of our characterizations of upper (resp. lower) operators. Before proceeding, we shall now recall some definitions and results which will be needed in the sequel. For any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$, we will denote the standard inner product of $\boldsymbol{x}$ and $\boldsymbol{y}$ by $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ and the Euclidean vector norm of $\boldsymbol{x} \in \mathbb{R}^d$ by $\|\boldsymbol{x}\| := \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$. Recall that a function $f : \Omega \to \mathbb{R}$ is called convex if for all $\boldsymbol{x}, \boldsymbol{y} \in \Omega, \lambda \in [0, 1]$: