Filomat 36:2 (2022), 695–701
https://doi.org/10.2298/FIL2202695A

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Optimal Estimates of Approximation Errors for Strongly Positive Linear Operators on Convex Polytopes

Osama Alabdali[a], **Allal Guessab**[b]

[a]*Department of Mathematics, College of Education for Pure Sciences, University Of Anbar, Ramadi, Iraq*
[b]*Laboratoire de Mathématiques et de leurs Applications, UMR CNRS 4152, E2S-Université de Pau et des Pays de l'Adour, 64000 Pau, France*

Abstract. In the present investigation, we introduce and study linear operators, which underestimate every strongly convex function. We call them, for brevity, sp−linear (approximation) operators. We will provide their sharp approximation errors. We show that the latter is bounded by the error approximation of the quadratic function. We use the centroidel Voronoi tessellations as a domain partition to construct best sp−linear operators. Finally, numerical examples are presented to illustrate the proposed method.

1. Introduction, motivation and theoretical justification

The problem of approximating a given function, that satisfies certain given conditions, is required in many applications. Generally, to get a better approximation of the function we try to approximate it in an appropriate candidate space, that satisfies all or a part of the given conditions. To describe our function approximation problem more precisely, let $f : \Omega \to \mathbb{R}$ be a given function, where Ω is a compact convex subset of $\mathbb{R}^d$. In some situations, we may know that f satisfies some type of convexity, we would like to use it in order to get a fairly good numerical integration of f. Our objective in this paper is to study linear operators, which underestimate all strongly convex functions. The notion of strong convexity takes its roots in the theory of numerical optimization. It is also of great use in mathematical economics, approximation theory and machine learning. Indeed, for function optimization methods, this concept of convexity has nice theoretical and practical properties. As we will see, one of the main advantage of using these special linear operators is that their approximation errors can be over or under-estimated in terms of the approximation error of the quadratic function.

To make things more concrete, let us start with a simple one dimensional motivating example. Assume given $\mu > 0$. One of the most successful strategy for approximating a given real μ-strongly convex function $f : [a, b] \to \mathbb{R}$ is first to choose a partition $P := \{x_0, x_1, \ldots, x_n\}$ of the interval $[a, b]$, such that $a = x_0 < x_1 < \ldots < x_n = b$, and then to approximate f using the first-order Taylor polynomial B_n about the midpoints of the subinterval $[x_i, x_{i+1}]$ such that B_n interpolates f and its first derivatives at

$$\frac{x_i + x_{i+1}}{2}, (i = 0, \ldots, n-1).$$

2020 *Mathematics Subject Classification*. Primary 65D30, 65D32, 41A63, 41A44, 41A80; Secondary 26B25, 26D15, 52A40.
Keywords. Multivariate approximate integration, convex functions, error estimates, Voronoi Diagram.
Received: 17 June 2019; Accepted: 20 July 2020
Communicated by Dragan S. Djordjević
Email addresses: alsltan_2006@yahoo.com (Osama Alabdali), allal.guessab@univ-pau.fr (Allal Guessab)

The operator B_n can be explicitly written in the form:

$$B_n[f](x) = \left(x - \frac{x_i + x_{i+1}}{2}\right) f'\left(\frac{x_i + x_{i+1}}{2}\right) + f\left(\frac{x_i + x_{i+1}}{2}\right), \quad (x \in [x_i, x_{i+1}]). \tag{1}$$

For the quadratic function $(.)^2$, the approximation error $E_n := f - B_n$ satisfies:

$$E_n\left[(.)^2\right](x) = \left(x - \frac{x_i + x_{i+1}}{2}\right)^2, \quad (x \in [x_i, x_{i+1}]). \tag{2}$$

In addition, for any μ -strongly convex function the following under estimate holds, for all $i = 0, \ldots, n-1$,

$$\frac{\mu}{2} E_n\left[(.)^2\right](x) \leq E_n[f](x), (x \in [x_i, x_{i+1}]). \tag{3}$$

Moreover, if the first derivative of f is Lipschitz continuous with papramter $L_i(f')$ in $[x_i, x_{i+1}]$, then the approximation error can be under and overestimated as:

$$\frac{\mu}{2} E_n\left[(.)^2\right](x) \leq E_n[f](x) \leq \frac{L_i(f')}{2} E_n\left[(.)^2\right](x). \tag{4}$$

Hence, the approximation error of this class of operators can always estimated in terms of the Lipschitz parameters of the first derivatives, the convexity parameter (of the strong convexity) and the error generated using the quadratic function. This provides a good starting point for discussion and further research. Indeed, the contributions of this paper are two-fold: first, our purpose is to extend this type of univariate results to the general multivariate variable case. More precisely, this paper deals with the problem of approximation of multivariable functions by using sp-linear operators. That is those which underestimate all strongly convex functions. A natural question is: can the approximation errors for such operators satisfy similar lower and upper bounds as given in (4) in the multivariate case?

Second, we construct a multivariate version of the operator of the form (1) in the case when the domain is a general polytope. Indeed, we use the centroidel Voronoi tessellations as a domain partition to construct best sp-linear operators. Finally, numerical examples are presented to illustrate the proposed method.

2. Optimal estimates of approximation errors

We will start in this section with some of the basic properties of strong convex functions. But first, we need to introduce some notations, which follow closely those of [1]. Let Ω be a nonempty and closed convex set of $\mathbb{R}^d$. We denote by $\|.\|$ the Euclidean norm in $\mathbb{R}^d$ and $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ the standard inner product of $x, y \in \mathbb{R}^d$. Let $C^{1,1}(\Omega)$ denote the set of all functions f which are continuously differentiable on Ω with Lipschitz continuous gradients, i.e., there exists $L[\nabla f]$ such that

$$\|\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})\| \leq L[\nabla f]\|\boldsymbol{x} - \boldsymbol{y}\|, (\boldsymbol{x}, \boldsymbol{y} \in \Omega).$$

A function f is said to be strongly convex with parameter $\mu > 0$ (written μ-strongly convex) if $g := f - \frac{\mu}{2}\|\cdot\|^2$ is convex. Let us denote the set of continuous convex functions $f : \Omega \to \mathbb{R}$ by $CC(\Omega)$, and the set of μ-strongly continuous convex functions on Ω by $SC_\mu(\Omega)$. We observe that $SC_\mu(\Omega)$ is obviously contained in $CC(\Omega)$.

Definition 2.1. *A linear operator $L : C(\Omega) \to C(\Omega)$ is called positive on $X \subset C(\Omega)$ if $L(f) \geq 0$ whenever $f \in X$.*

The following Lemma provides an easy but important property, it will be used extensively throughout the reminder of this work. Indeed, it gives a necessary and sufficient condition for the positivity of linear operators on $CC(\Omega)$. It is shown that in order to prove the positivity of a linear operator on $CC(\Omega)$, it suffices to verify that the property is satisfied by the elements in the subset $SC_\mu(\Omega)$ for a given fixed strong convexity parameter μ.

Lemma 2.2. *Given μ a arbitrary fixed positive number. If $E : C(\Omega) \to C(\Omega)$ is a linear operator, then E is positive on $CC(\Omega)$ if and only if E is positive on $SC_\mu(\Omega)$.*

Proof. The direct implication is easy, since $SC_\mu(\Omega) \subset CC(\Omega)$. For the other way implication assume that E is positive on $SC_\mu(\Omega)$. Let $\varepsilon > 0$ and let f be a convex function. Define

$$g := f + \frac{\varepsilon}{2}\|\cdot\|^2.$$

Multiplying by $\frac{\mu}{\varepsilon}$ and rearranging, we obtain

$$\frac{\mu}{\varepsilon} f = \frac{\mu}{\varepsilon} g - \frac{\mu}{2}\|\cdot\|^2$$

and since $\frac{\mu}{\varepsilon} f$ is convex, then $\frac{\mu}{\varepsilon} g$ is μ-strongly convex. But E is positive on $SC_\mu(\Omega)$ then

$$E\left(\frac{\mu}{\varepsilon} g\right) \geq 0.$$

Thus by homogeneity of E, it is immediate that

$$E(g) \geq 0.$$

It is easily derived from the linearity of E that

$$E(f) \geq -\frac{\varepsilon}{2} E\left(\|\cdot\|^2\right).$$

Since this inequality holds for all $\varepsilon > 0$, then by letting $\varepsilon \downarrow 0$, we get

$$E(f) \geq 0.$$

This yields the desired result and completes the proof of Lemma 2.2. □

3. Characterizations of sp-linear operators

In the following, we say that a linear operator $A : X \subset C^1(\Omega) \to C(\Omega)$ underestimates the identity operator on X if, for all $f \in X$ and $x \in \Omega, A[f](x) \leq f(x)$. We observe that a linear operator A satisfies this property if and only if the approximation error $I - A$ is positive on X. The characterization of positive linear operators given by Lemma 2.2 allows us to provide an error characterization estimate of these latter.

Theorem 3.1. *Let $\mu > 0$ and let $A : C^1(\Omega) \to C(\Omega)$ be a linear operator. Then, A underestimates the identity operator on SC_μ if and only if for every $f \in C^{1,1}(\Omega)$, the approximation error $E := I - A$ satisfies*

$$|E[f](x)| \leq \frac{L(\nabla f)}{2}\left(E\left[\|\cdot\|^2\right](x)\right), \quad (x \in \Omega), \tag{5}$$

where $L(\nabla f)$ denotes the Lipschitz constant of the gradient of f.

Proof. This result is essentially based on [2, Theorem 2.3] proved in the case of the classical convexity. Assume that A underestimates the identity operator on SC_μ. Then, it follows that E is positive on $SC_\mu(\Omega)$. Therefore, by an application of Lemma 2.2 we deduce that E is also positive on $CC(\Omega)$. Now [2, Theorem 2.3] shows that (5) is satisfied. Conversely, assume that (5) holds. Then, again a simple application of [2, Theorem 2.3] implies that E is positive on $CC(\Omega)$. Thus, Lemma 2.2 yields E is also positive on $SC(\Omega)$. □

The following Corollary, which is in part is a consequence of Theorem 3.1 improves estimate (5) when the strong convexity is also imposed on the function, which we want to approximate.

Corollary 3.2. *Let $\mu > 0$. Let $A : C^1(\Omega) \to C(\Omega)$ be a linear operator. Assume that A underestimates the identity operator on $SC_\mu(\Omega)$. Then for every $f \in SC_\mu(\Omega) \cap C^{1,1}(\Omega)$, the approximation operator $E := I - A$ satisfies*

$$\frac{\mu}{2}\left(E\left[\|\cdot\|^2\right](x)\right) \le E[f](x) \le \frac{L(\nabla f)}{2}\left(E\left[\|\cdot\|^2\right](x)\right), \quad (x \in \Omega). \tag{6}$$

Proof. By Theorem 3.1 it remains to show that the lower bound holds. Assume that A underestimates the identity operator on $SC(\Omega)$ and let us fix $f \in SC_\mu(\Omega)$. By Lemma 2.2 the approximation error $I - A$ is positive on $CC(\Omega)$. Since $g = f - \frac{\mu}{2}\|\cdot\|^2$ is convex, then Lemma 2.2 applied to $I - A$ implies

$$f - \frac{\mu}{2}\|\cdot\|^2 \ge A[f] - \frac{\mu}{2}A\left[\|\cdot\|^2\right]$$

and therefore, we get after some manipulations

$$\frac{\mu}{2}\left(\|\cdot\|^2 - A\left[\|\cdot\|^2\right]\right) \le f - A[f].$$

This yields the desired result. The case of equality can be verified by a simple calculation. □

The over and underestimates (6) tell us that if the approximation error associated to the quadratic function is small, then we are confident that those of strongly convex functions is also small.

4. Applications to the weighted averaging approximation

Let us assume that $X_m = \{x_i\}_{i=0}^m \subset \Omega \subset \mathbb{R}^d$, with $\Omega = \operatorname{conv}(X_m)$. We are interested in approximating an unknown function $f : \Omega \to \mathbb{R}$ from given function values and its gradient $f(y_0), \nabla f(y_0), \ldots, f(y_n), \nabla f(y_n)$ where $Y_n := \{y_i\}_{i=0}^n \subset \Omega$. Consider the weighted averaging operator (WAO)

$$B_n[f](x) = \sum_{i=0}^{n} \lambda_i(x)\left(f(y_i) + \langle \nabla f(y_i), x - y_i\rangle\right). \tag{7}$$

This means that we impose the normalizing condition on the system of functions $\lambda := \{\lambda_i\}_{i=0}^n$ for all $x \in \Omega$

$$\begin{aligned} &\lambda_i(x) \ge 0, \quad i = 0, \ldots, n \\ &\sum_{i=0}^{n} \lambda_i(x) = 1. \end{aligned} \tag{8}$$

We will describe here our approach for constructing a set of functions $\lambda := \{\lambda_i\}_{i=0}^n$, that yields a good approximation operator B_n. To that end, let us introduce the notion of the Voronoi tessellations for a set of distinct points $y_0, \ldots, y_n \in \Omega$. The Voronoi sets generated by these points are defined for $i = 0, \ldots, n$, by

$$\Omega_i = \left\{x \in \Omega : \left\|x - y_i\right\| \le \left\|x - y_j\right\|, j = 0, \ldots, n, j \ne i\right\}.$$

The domain partition $\Omega_0, \ldots, \Omega_n$ of Ω is called a Voronoi tessellation. It is said to be centroidal if

$$\boldsymbol{y}_i = \frac{1}{|\Omega_i|}\int_{\Omega_i} \boldsymbol{x} d\boldsymbol{x}.$$

This means that the centre of gravity will always be the same as the generator of any Voronoi region in a centroidal Voronoi tessellation (CVT).
For any function $f \in C^{1,1}(\Omega)$, we define the associated error to B_n as

$$E_n[f](\boldsymbol{x}) := E_n[f, \boldsymbol{\lambda}](\boldsymbol{x}) = f(\boldsymbol{x}) - B_n[f](\boldsymbol{x}), (\boldsymbol{x} \in \Omega). \tag{9}$$

The following lemma provides a simple explicit expression of the approximation error $E_n\left[\|\cdot\|^2\right]$.

Lemma 4.1. *The approximation error* $E_n\left[\|.\|^2\right]$ *has the convenient form:*

$$E_n\left[\|\cdot\|^2\right](x) = \sum_{i=0}^{n} \lambda_i(x) \left\|x - y_i\right\|^2. \tag{10}$$

Proof. If $f(\boldsymbol{x}) = \|\boldsymbol{x}\|^2$, we easily get

$$f\left(\boldsymbol{y}_i\right) + \left\langle \nabla f\left(\boldsymbol{y}_i\right), \boldsymbol{x} - \boldsymbol{y}_i\right\rangle = \|\boldsymbol{x}\|^2 - \left\|\boldsymbol{x} - \boldsymbol{y}_i\right\|^2.$$

Hence, multiplying on each side by λ_i, summing up with respect to i from 0 to n, we get desired result after arrangement. □

We also have:

Lemma 4.2. *Let* $\mu > 0$*. Then, the operator* B_n *underestimates the identity operator on* SC_μ*. Moreover, for every* $f \in SC_\mu(\Omega)$*, it holds*

$$\frac{\mu}{2} E_n\left[\|\cdot\|^2\right](x) = \frac{\mu}{2} \sum_{i=0}^{n} \lambda_i(x) \left\|x - y_i\right\|^2 \leq E_n[f](x), \quad (x \in \Omega). \tag{11}$$

Proof. The equality sign in (11) has already proved in Lemma 4.1. Let us fix f a μ-strongly convex function. By the Jensen-convexity for μ-strongly convex functions, see [4] we get

$$f(x) \geq f\left(y_i\right) + \left\langle \nabla f\left(y_i\right), x - y_i\right\rangle + \frac{\mu}{2} \left\|x - y_i\right\|^2.$$

Hence, multiplying on each side by λ_i, summing up with respect to i from 0 to n and rearranging, we get the required result. □

The following Lemma gives an upper bound for the absolute value of the approximation error.

Lemma 4.3. *For every* $f \in C^{1,1}(\Omega)$*, it holds*

$$\left|E_n[f](x)\right| \leq \frac{L(\nabla f)}{2} \sum_{i=0}^{n} \lambda_i(x) \left\|x - y_i\right\|^2, \quad (x \in \Omega). \tag{12}$$

Proof. This Lemma is an immediate consequence of Theorem 3.1 and Lemma 4.1 □

The following result, which shows that the approximation error is dominated by the approximation error of the quadratic function, will be important for the applications in the next section.

Theorem 4.4. *Let* $\mu > 0$*. Then, for every function* $f \in SC_\mu(\Omega) \cap C_\mu^{1,1}(\Omega)$ *and any* $x \in \Omega$*, it holds:*

$$\frac{\mu}{2} E_n\left[\|\cdot\|^2\right](\boldsymbol{x}) \leq E_n[f](\boldsymbol{x}) \leq \frac{L(\nabla f)}{2} E_n\left[\|\cdot\|^2\right](\boldsymbol{x}). \tag{13}$$

Proof. This is an immediate consequence of Lemmas 4.2, 4.3 and Theorem 3.1. The case of equality is easily verified. □

From the over and underestimates (13), it is important to find a good approximate for the quadratic function. By Lemma 4.1, the quality of our approximation depends critically not only on the partition of unity λ but also on the interpolation points $y_0, \ldots, y_n$. Thus, a natural question now arises: how small can $E_n\left[\|.\|^2\right]$ be over all possible choices for the partition of unity λ and the points $y_0, \ldots, y_n$?

The answer for any $p \geq 1$ can be found in [3, Theorem 2.11].

Theorem 4.5. *For $p \in [1, \infty)$ there exists an operator B_n of the form (8) which is optimal with respect to the L^p norm. Denoting by $\Omega_0, \ldots, \Omega_n$ the CVT by $y_0, \ldots, y_n$, we have $\lambda_i = 1_{\Omega_i}$ almost everywhere on Ω for $i = 0, \ldots, n$ and, in addition, the equations*

$$\boldsymbol{y}_i \int_{\Omega_i} \left\|\boldsymbol{x} - \boldsymbol{y}_i\right\|^{2p-2} d\boldsymbol{x} = \int_{\Omega_i} \boldsymbol{x} \left\|\boldsymbol{x} - \boldsymbol{y}_i\right\|^{2p-2} d\boldsymbol{x}, (i = 0, \ldots, n)$$

are satisfied.

5. Numerical experiments

We have seen above (Section 4) that an interesting weighted averaging approximation operator would be to simply construct a centroidal voronoi diagram generated by a set of random points

$$Y_n = \left\{\boldsymbol{y}_i\right\}_{i=0}^{N}.$$

In order to give numerical illustrations of the performance of the implementation of our approach, we apply the method to the reconstruction of two test functions $g_k, k = 1, 2$, when the domain Ω is the unit square $[0, 1] \times [0, 1]$, and the function g_k exhibits the following features: it is sufficiently regular, it is strongly convex, and can be evaluated at any point of the domain. For each of the two test functions g_k, we take N scattered points $\{y_i\}_{i=1}^{N}$, which are the generators of a centroidal Voronoi tessellation on Ω, and construct the operator $B_N[g_k]$. We then determine the mean square error (*MSE*) by evaluating

$$\sqrt{\sum_{i=1}^{N} \frac{\left(f_k\left(\boldsymbol{y}_i\right) - B_N[g_k]\left(\boldsymbol{y}_i\right)\right)^2}{N}}.$$

at N

Example 5.1. *We take:*

$$g_1(x, y) = 0.3\left((x - 0.5)^2 + (y + 0.6)^2\right) + 0.3\exp\left((x - 0.3)^2 + (y - 0.3)^2\right)$$

and $D := [0, 1] \times [0, 1]$

Function	*Number of scatter data*	*MSE*
$g_1(x, y)$	*50*	5.8×10^{-3}
	250	6.8005×10^{-4}
	1300	5.6326×10^{-5}

Example 5.2. *Here, we take the following strongly convex function*

$$g_2(x, y) = 0.4\left((x - 0.5)^2 + (y + 0.6)^2\right).$$

Function	*Number of scatter data*	*MSE*
$g_2(x, y)$	*50*	1.8×10^{-3}
	300	1.9338×10^{-4}
	700	6.4765×10^{-5}

In summary, we obtain encouraging results with few random points on the domain.

Acknowledgement

This paper is dedicated to Professor Gradimir V. Milovanovic on the Occasion of his 70 th Anniversary.

References

[1] Alabdali, O., Guessab, A., On the approximation of strongly convex function by an upper or lower operator, Appl. Math.Comput., 247(2014), 1129–1138.
[2] Guessab, A., Approximations of differentiable convex functions on arbitrary convex polytopes, Appl. Math. Comput., 240(2014), 326–338.
[3] Guessab, A., Schmeisser, G., Construction of positive definite cubature formulae and approximation of functions viaVoronoi tessellations, Adv. Comput. Math., 32(2010), 25–41.
[4] J.B. Hiriart-Urruty, C. Lemaréchal, Fundamentals of Convex Analysis, Springer, Berlin (2001).