INTERDISCIPLINARY MATHEMATICS INSTITUTE

2014:05

Incremental Greedy Algorithm and its Applications in Numerical Integration

V. Temlyakov

IMI
PREPRINT SERIES

COLLEGE OF ARTS AND SCIENCES
UNIVERSITY OF SOUTH CAROLINA

Incremental Greedy Algorithm and its applications in numerical integration*

V. Temlyakov [†]

September 17, 2014

Abstract

Applications of the Incremental Algorithm, which was developed in the theory of greedy algorithms in Banach spaces, to approximation and numerical integration are discussed. In particular, it is shown that the Incremental Algorithm provides an efficient way for deterministic construction of cubature formulas with equal weights, which give good rate of error decay for a wide variety of function classes.

1 Introduction

The paper provides some progress in the fundamental problem of algorithmic construction of good methods of approximation and numerical integration. Numerical integration seeks good ways of approximating an integral

$$\int_\Omega f(x)d\mu$$

by an expression of the form

$$\Lambda_m(f,\xi) := \sum_{j=1}^m \lambda_j f(\xi^j), \quad \xi = (\xi^1,\dots,\xi^m), \quad \xi^j \in \Omega, \quad j = 1,\dots,m. \tag{1.1}$$

Math Subject Classifications. primary: 41A65; secondary: 42A10, 46B20.

[†]University of South Carolina, USA, and Steklov Institute of Mathematics, Russia.
Research was supported by NSF grant DMS-1160841

It is clear that we must assume that f is integrable and defined at the points $\xi^1, \dots, \xi^m$. The expression (1.1) is called a cubature formula (Λ, ξ) (if $\Omega \subset \mathbb{R}^d$, $d \geq 2$) or a quadrature formula (Λ, ξ) (if $\Omega \subset \mathbb{R}$) with knots $\xi = (\xi^1, \dots, \xi^m)$ and weights $\Lambda = (\lambda_1, \dots, \lambda_m)$. For a function class W we introduce a concept of error of the cubature formula $\Lambda_m(\cdot, \xi)$ by

$$\Lambda_m(W, \xi) := \sup_{f \in W} \left| \int_\Omega f d\mu - \Lambda_m(f, \xi) \right|. \tag{1.2}$$

There are many different ways to construct good deterministic cubature formulas beginning with heuristic guess of good knots for a specific class and ending with finding a good cubature formula as a solution (approximate solution) of the optimization problem

$$\inf_{\xi^1, \dots, \xi^m; \lambda_1, \dots, \lambda_m} \Lambda_m(W, \xi).$$

Clearly, the way of solving the above optimization problem is the preferable one. However, in many cases this problem is very hard (see a discussion in [10]). It was observed in [9] that greedy-type algorithms provide an efficient way for deterministic constructions of good cubature formulas for a wide variety of function classes. This paper is a follow up to [9]. In this paper we discuss in detail a greedy-type algorithm – Incremental Algorithm – that was not discussed in [9]. The main advantage of the Incremental Algorithm over the greedy-type algorithms considered in [9] is that it provides better control of weights of the cubature formula and gives the same rate of decay of the integration error.

We remind some notations from the theory of greedy approximation in Banach spaces. The reader can find a systematic presentation of this theory in [11], Chapter 6. Let X be a Banach space with norm $\|\cdot\|$. We say that a set of elements (functions) $\mathcal{D}$ from X is a dictionary if each $g \in \mathcal{D}$ has norm less than or equal to one ($\|g\| \leq 1$) and the closure of $\operatorname{span} \mathcal{D}$ coincides with X. We note that in [8] we required in the definition of a dictionary normalization of its elements ($\|g\| = 1$). However, it is pointed out in [10] that it is easy to check that the arguments from [8] work under assumption $\|g\| \leq 1$ instead of $\|g\| = 1$. In applications it is more convenient for us to have an assumption $\|g\| \leq 1$ than normalization of a dictionary.

For an element $f \in X$ we denote by F_f a norming (peak) functional for f:

$$\|F_f\| = 1, \qquad F_f(f) = \|f\|.$$

The existence of such a functional is guaranteed by the Hahn-Banach theorem.

We proceed to the Incremental Greedy Algorithm (see [10] and [11], Chapter 6). Let $\epsilon = \{\epsilon_n\}_{n=1}^{\infty}$, $\epsilon_n > 0$, $n = 1, 2, \dots$.

Incremental Algorithm with schedule ϵ (IA(ϵ)). Denote $f_0^{i,\epsilon} := f$ and $G_0^{i,\epsilon} := 0$. Then, for each $m \geq 1$ we have the following inductive definition.

(1) $\varphi_m^{i,\epsilon} \in \mathcal{D}$ is any element satisfying

$$F_{f_{m-1}^{i,\epsilon}}(\varphi_m^{i,\epsilon} - f) \geq -\epsilon_m.$$

(2) Define

$$G_m^{i,\epsilon} := (1 - 1/m)G_{m-1}^{i,\epsilon} + \varphi_m^{i,\epsilon}/m.$$

(3) Let

$$f_m^{i,\epsilon} := f - G_m^{i,\epsilon}.$$

We show how the Incremental Algorithm can be used in approximation and numerical integration. We begin with a discussion of the approximation problem. A detailed discussion, including historical remarks, is presented in Section 2. For simplicity, we illustrate how the Incremental Algorithm works in approximation of univariate trigonometric polynomials.

An expression

$$\sum_{j=1}^{m} c_j g_j, \quad g_j \in \mathcal{D}, \quad c_j \in \mathbb{R}, \quad j = 1, \dots, m$$

is called m-term polynomial with respect to $\mathcal{D}$. The concept of best m-term approximation with respect to $\mathcal{D}$

$$\sigma_m(f, \mathcal{D})_X := \inf_{\{c_j\}, \{g_j \in \mathcal{D}\}} \|f - \sum_{j=1}^{m} c_j g_j\|_X$$

plays an important role in our consideration.

By $\mathcal{RT}(N)$ we denote the set of real 1-periodic trigonometric polynomials of order N. For a real trigonometric polynomial denote

$$\|a_0 + \sum_{k=1}^{N}(a_k \cos k2\pi x + b_k \sin k2\pi x)\|_A := |a_0| + \sum_{k=1}^{N}(|a_k| + |b_k|).$$

We formulate here a result from [10].

Theorem 1.1. *There exists a constructive method $A(N,m)$ such that for any $t \in \mathcal{RT}(N)$ it provides a m-term trigonometric polynomial $A(N,m)(t)$ with the following approximation property*

$$\|t - A(N,m)(t)\|_\infty \le Cm^{-1/2}(\ln(1+N/m))^{1/2}\|t\|_A$$

with an absolute constant C.

An advantage of the IA(ϵ) over other greedy-type algorithms is that the IA(ϵ) gives precise control of the coefficients of the approximant. For all approximants $G_m^{i,\epsilon}$ we have the property $\|G_m^{i,\epsilon}\|_A = 1$. Moreover, we know that all nonzero coefficients of the approximant have the form a/m where a is a natural number. In Section 2 we prove the following result.

Theorem 1.2. *For any $t \in \mathcal{RT}(N)$ the IA(ϵ) applied to $f := t/\|t\|_A$ provides after m iterations a m-term trigonometric polynomial $G_m(t) := G_m^{i,\epsilon}(f)\|t\|_A$ with the following approximation property*

$$\|t - G_m(t)\|_\infty \le Cm^{-1/2}(\ln N)^{1/2}\|t\|_A, \quad \|G_m(t)\|_A = \|t\|_A,$$

with an absolute constant C.

We note that the implementation of the IA(ϵ) depends on the dictionary and the ambient space X. The IA(ϵ) from Theorem 1.2 acts with respect to the real trigonometric system $1, \cos 2\pi x, \sin 2\pi x, \dots, \cos N2\pi x, \sin N2\pi x$ in the space $X = L_p$ with $p \asymp \ln N$.

We now proceed to results from Section 3 on numerical integration. As in [9] we define a set $\mathcal{K}_q$ of kernels possessing the following properties. Let $K(x,y)$ be a measurable function on $\Omega_x \times \Omega_y$. We assume that for any $x \in \Omega_x$ $K(x,\cdot) \in L_q(\Omega_y)$, for any $y \in \Omega_y$ the $K(\cdot,y)$ is integrable over Ω_x and $\int_{\Omega_x} K(x,\cdot)dx \in L_q(\Omega_y)$.

For a kernel $K \in \mathcal{K}_{p'}$ we define the class

$$W_p^K := \{f : f = \int_{\Omega_y} K(x,y)\varphi(y)dy, \quad \|\varphi\|_{L_p(\Omega_y)} \le 1\}.$$

Then each $f \in W_p^K$ is integrable on Ω_x (by Fubini's theorem) and defined at each point of Ω_x. We denote for convenience

$$J(y) := J_K(y) := \int_{\Omega_x} K(x,y)dx.$$

Consider a dictionary

$$\mathcal{D} := \{K(x,\cdot), x \in \Omega_x\}$$

and define a Banach space $X(K,p')$ as the $L_{p'}(\Omega_y)$-closure of span of $\mathcal{D}$. In Section 3 the following theorem is proved.

Theorem 1.3. *Let W^K_p be a class of functions defined above. Assume that $K \in \mathcal{K}_{p'}$ satisfies the condition*

$$\|K(x,\cdot)\|_{L_{p'}(\Omega_y)} \le 1, \quad x \in \Omega_x, \quad |\Omega_x| = 1$$

and $J_K \in X(K,p')$. Then for any m there exists (provided by the Incremental Algorithm) a cubature formula $\Lambda_m(\cdot,\xi)$ with $\lambda_\mu = 1/m$, $\mu = 1,2,\dots,m$, and

$$\Lambda_m(W^K_p,\xi) \le C(p-1)^{-1/2}m^{-1/2}, \quad 1 < p \le 2.$$

Theorem 1.3 provides a constructive way of finding for a wide variety of classes W^K_p cubature formulas that give the error bound similar to that of the Monter Carlo method. We stress that in Theorem 1.3 we do not assume any smoothness of the kernel $K(x,y)$.

2 Approximation by the Incremental Algorithm

First, we discuss the known Theorem 1.1 from the Introduction. Proof of Theorem 1.1 is based on a greedy-type algorithm – the Weak Chebyshev Greedy Algorithm. We now describe it. Let $\tau := \{t_k\}_{k=1}^\infty$ be a given sequence of nonnegative numbers $t_k \le 1$, $k = 1,\dots$. We define (see [8]) the Weak Chebyshev Greedy Algorithm (WCGA) that is a generalization for Banach spaces of Weak Orthogonal Greedy Algorithm defined and studied in [7] (see also [11]).

Weak Chebyshev Greedy Algorithm (WCGA). We define $f^c_0 := f^{c,\tau}_0 := f$. Then for each $m \ge 1$ we inductively define

1). $\varphi^c_m := \varphi^{c,\tau}_m \in \mathcal{D}$ is any element satisfying

$$|F_{f^c_{m-1}}(\varphi^c_m)| \ge t_m \sup_{g\in\mathcal{D}} |F_{f^c_{m-1}}(g)|.$$

2). Define

$$\Phi_m := \Phi^\tau_m := \operatorname{span}\{\varphi^c_j\}_{j=1}^m,$$

and define $G_m^c := G_m^{c,\tau}$ to be the best approximant to f from Φ_m.

3). Denote

$$f_m^c := f_m^{c,\tau} := f - G_m^c.$$

The term "weak" in this definition means that at the step 1) we do not shoot for the optimal element of the dictionary, which realizes the corresponding supremum, but are satisfied with weaker property than being optimal. The obvious reason for this is that we do not know in general that the optimal one exists. Another, practical reason is that the weaker the assumption the easier to satisfy it and, therefore, easier to realize in practice.

We consider here approximation in uniformly smooth Banach spaces. For a Banach space X we define the modulus of smoothness

$$\rho(u) := \sup_{\|x\|=\|y\|=1} (\frac{1}{2}(\|x+uy\| + \|x-uy\|) - 1).$$

The uniformly smooth Banach space is the one with the property

$$\lim_{u\to 0} \rho(u)/u = 0.$$

It is well known (see for instance [3], Lemma B.1) that in the case $X = L_p$, $1 \le p < \infty$ we have

$$\rho(u) \le \begin{cases} u^p/p & \text{if} \quad 1 \le p \le 2, \\ (p-1)u^2/2 & \text{if} \quad 2 \le p < \infty. \end{cases} \tag{2.1}$$

Denote by $A_1(\mathcal{D}) := A_1(\mathcal{D}, X)$ the closure in X of the convex hull of $\mathcal{D}$. The following theorem from [8] gives the rate of convergence of the WCGA for f in $A_1(\mathcal{D})$.

Theorem 2.1. *Let X be a uniformly smooth Banach space with the modulus of smoothness $\rho(u) \le \gamma u^q$, $1 < q \le 2$. Then for $t \in (0,1]$ we have for any $f \in A_1(\mathcal{D})$ that*

$$\|f - G_m^{c,\tau}(f,\mathcal{D})\| \le C(q,\gamma)(1 + mt^p)^{-1/p}, \quad p := \frac{q}{q-1},$$

with a constant $C(q,\gamma)$ which may depend only on q and γ.

In [10] we demonstrated the power of the WCGA in classical areas of harmonic analysis. The problem concerns the trigonometric m-term approximation in the uniform norm. The first result that indicated an advantage of m-term approximation with respect to the real trigonometric system $\mathcal{RT}$ over approximation by trigonometric polynomials of order m is due to Ismagilov [5]

$$\sigma_m(|\sin 2\pi x|, \mathcal{RT})_\infty \le C_\epsilon m^{-6/5+\epsilon}, \quad \text{for any} \quad \epsilon > 0. \tag{2.2}$$

Maiorov [6] improved the estimate (2.2):

$$\sigma_m(|\sin 2\pi x|, \mathcal{RT})_\infty \asymp m^{-3/2}. \tag{2.3}$$

Both R.S. Ismagilov [5] and V.E. Maiorov [6] used constructive methods to get their estimates (2.2) and (2.3). V.E. Maiorov [6] applied a number theoretical method based on Gaussian sums. The key point of that technique can be formulated in terms of best m-term approximation of trigonometric polynomials. Let as above $\mathcal{RT}(N)$ be the subspace of real trigonometric polynomials of order N. Using the Gaussian sums one can prove (constructively) the estimate

$$\sigma_m(t, \mathcal{RT})_\infty \le CN^{3/2}m^{-1}\|t\|_1, \quad t \in \mathcal{RT}(N). \tag{2.4}$$

Denote as above

$$\|a_0 + \sum_{k=1}^N (a_k \cos k2\pi x + b_k \sin k2\pi x)\|_A := |a_0| + \sum_{k=1}^N (|a_k| + |b_k|).$$

We note that by simple inequality

$$\|t\|_A \le (2N+1)\|t\|_1, \quad t \in \mathcal{RT}(N),$$

the estimate (2.4) follows from the estimate

$$\sigma_m(t, \mathcal{RT})_\infty \le C(N^{1/2}/m)\|t\|_A, \quad t \in \mathcal{RT}(N). \tag{2.5}$$

Thus (2.5) is stronger than (2.4). The following estimate was proved in [1]

$$\sigma_m(t, \mathcal{RT})_\infty \le Cm^{-1/2}(\ln(1+N/m))^{1/2}\|t\|_A, \quad t \in \mathcal{RT}(N). \tag{2.6}$$

In a way (2.6) is much stronger than (2.5) and (2.4). The proof of (2.6) from [1] is not constructive. The estimate (2.6) has been proved in [1] with the help of a nonconstructive theorem of Gluskin [4]. In [10] we gave a constructive proof of (2.6). The key ingredient of that proof is the WCGA. In the paper [2] we already pointed out that the WCGA provides a constructive proof of the estimate

$$\sigma_m(f, \mathcal{RT})_p \le C(p) m^{-1/2} \|f\|_A, \quad p \in [2, \infty). \tag{2.7}$$

The known proofs (before [2]) of (2.7) were nonconstructive (see discussion in [2], Section 5). Thus, the WCGA provides a way of building a good m-term approximant. However, the step 2) of the WCGA makes it difficult to control the coefficients of the approximant – they are obtained through the Chebyshev projection of f onto Φ_m. This motivates us to consider the IA(ϵ) which gives explicit coefficients of the approximant.

Second, we proceed to a discussion and proof of Theorem 1.2. In order to be able to run the IA(ϵ) for all iterations we need existence of an element $\varphi_m^{i,\epsilon} \in \mathcal{D}$ at the step (1) of the algorithm for all m. It is clear that the following condition guarantees such existence.

Condition B. We say that for a given dictionary $\mathcal{D}$ an element f satisfies Condition B if for any $F \in X^*$ we have

$$F(f) \le \sup_{g \in \mathcal{D}} F(g).$$

It is well known (see, for instance, [11], p. 343) that any $f \in A_1(\mathcal{D})$ satisfies Condition B. For completeness we give this simple argument here. Take any $f \in A_1(\mathcal{D})$. Then for any $\epsilon > 0$ there exist $g_1^\epsilon, \dots, g_N^\epsilon \in \mathcal{D}$ and numbers $a_1^\epsilon, \dots, a_N^\epsilon$ such that $a_i^\epsilon > 0$, $a_1^\epsilon + \dots + a_N^\epsilon = 1$ and

$$\|f - \sum_{i=1}^N a_i^\epsilon g_i^\epsilon\| \le \epsilon.$$

Thus

$$F(f) \le \|F\|\epsilon + F(\sum_{i=1}^N a_i^\epsilon g_i^\epsilon) \le \epsilon\|F\| + \sup_{g \in \mathcal{D}} F(g)$$

which proves Condition B.

We note that Condition B is equivalent to the property $f \in A_1(\mathcal{D})$. Indeed, as we showed above, the property $f \in A_1(\mathcal{D})$ implies Condition B. Let us show that Condition B implies that $f \in A_1(\mathcal{D})$. Assuming the contrary

$f \notin A_1(\mathcal{D})$ by separation theorem for convex bodies we find $F \in X^*$ such that

$$F(f) > \sup_{\phi \in A_1(\mathcal{D})} F(\phi) \geq \sup_{g \in \mathcal{D}} F(g)$$

which contradicts Condition B.

We formulate results on the AI(ϵ) in terms of Condition B because in the application from Section 3 it is easy to check Condition B.

Theorem 2.2. *Let X be a uniformly smooth Banach space with modulus of smoothness $\rho(u) \leq \gamma u^q$, $1 < q \leq 2$. Define*

$$\epsilon_n := \beta \gamma^{1/q} n^{-1/p}, \qquad p = \frac{q}{q-1}, \quad n = 1, 2, \dots.$$

Then, for any f satisfying Condition B we have

$$\|f_m^{i,\epsilon}\| \leq C(\beta) \gamma^{1/q} m^{-1/p}, \qquad m = 1, 2 \dots.$$

In the case $f \in A_1(\mathcal{D})$ this theorem is proved in [10] (see also [11], Chapter 6). As we mentioned above Condition B is equivalent to $f \in A_1(\mathcal{D})$.

We now give some applications of Theorem 2.2 in construction of special polynomials. We begin with a general result.

Theorem 2.3. *Let X be a uniformly smooth Banach space with modulus of smoothness $\rho(u) \leq \gamma u^q$, $1 < q \leq 2$. For any n elements $\varphi_1, \varphi_2, \dots, \varphi_n$, $\|\varphi_j\| \leq 1$, $j = 1, \dots, n$, there exists a subset $\Lambda \subset [1, n]$ of cardinality $|\Lambda| \leq m < n$ and natural numbers a_j, $j \in \Lambda$ such that*

$$\|\frac{1}{n} \sum_{j=1}^{n} \varphi_j - \sum_{j \in \Lambda} \frac{a_j}{m} \varphi_j \|_X \leq C \gamma^{1/q} m^{1/q - 1}, \qquad \sum_{j \in \Lambda} a_j = m.$$

Proof. For a given set $\varphi_1, \varphi_2, \dots, \varphi_n$ consider a new Banach space $X_n := \operatorname{span}(\varphi_1, \varphi_2, \dots, \varphi_n)$ with norm $\|\cdot\|_X$. In the space X_n consider the dictionary $\mathcal{D}_n := \{\varphi_j\}_{j=1}^n$. Then the space X_n is a uniformly smooth Banach space with modulus of smoothness $\rho(u) \leq \gamma u^q$, $1 < q \leq 2$ and $f := \frac{1}{n} \sum_{j=1}^n \varphi_j \in A_1(\mathcal{D}_n)$. Applying the AI($\epsilon$) to f with respect to $\mathcal{D}_n$ we obtain by Theorem 2.2 after m iterations

$$\|f - \sum_{k=1}^{m} \frac{1}{m} \varphi_{j_k} \|_X \leq C \gamma^{1/q} m^{1/q - 1},$$

where φ_{j_k} is obtained at the kth iteration of the AI(ϵ). Clearly, $\sum_{k=1}^m \frac{1}{m} \varphi_{j_k}$ can be written in the form $\sum_{j \in \Lambda} \frac{a_j}{m} \varphi_j$ with $|\Lambda| \leq m$. $\square$

Corollary 2.1. *Let $m \in \mathbb{N}$ and $n = 2m$. For any n trigonometric polynomials $\varphi_j \in \mathcal{RT}(N)$, $\|\varphi_j\|_\infty \le 1$, $j = 1, \dots, n$ with $N \le n^b$ there exist a set Λ and natural numbers a_j, $j \in \Lambda$, such that $|\Lambda| \le m$, $\sum_{j\in\Lambda} a_j = m$ and*

$$\|\frac{1}{n}\sum_{j=1}^{n}\varphi_j - \sum_{j\in\Lambda}\frac{a_j}{m}\varphi_j\|_\infty \le C(b)(\ln m)^{1/2}m^{-1/2}. \tag{2.8}$$

Proof. First, we apply Theorem 2.3 with $X = L_p$, $2 \le p < \infty$. Using (2.1) we get

$$\|\frac{1}{n}\sum_{j=1}^{n}\varphi_j - \sum_{j\in\Lambda(p)}\frac{a_j(p)}{m}\varphi_j\|_p \le Cp^{1/2}m^{-1/2}, \qquad \sum_{j\in\Lambda(p)} a_j(p) = m, \tag{2.9}$$

with $|\Lambda(p)| \le m$.

Second, by the Nikol'skii inequality we obtain from (2.9)

$$\|\frac{1}{n}\sum_{j=1}^{n}\varphi_j - \sum_{j\in\Lambda(p)}\frac{a_j(p)}{m}\varphi_j\|_\infty$$

$$\le CN^{1/p}\|\frac{1}{n}\sum_{j=1}^{n}\varphi_j - \sum_{j\in\Lambda(p)}\frac{a_j(p)}{m}\varphi_j\|_p \le Cp^{1/2}N^{1/p}m^{-1/2}.$$

Choosing $p \asymp \ln N \asymp \ln m$ we obtain (2.8). $\square$

We note that Corollary 2.1 provides a construction of analogs of the Rudin-Shapiro polynomials in a much more general situation than in the case of the Rudin-Shapiro polynomials, albeit with a little bit weaker bound, which contains an extra $(\ln m)^{1/2}$ factor.

Proof of Theorem 1.2. It is clear that it is sufficient to prove Theorem 1.2 for $t \in \mathcal{RT}(N)$ with $\|t\|_A = 1$. Then $t \in A_1(\mathcal{RT}(N), L_p)$ for all $p \in [2, \infty)$. Now, applying Theorem 2.3 and using its proof with $X = L_p$, $\varphi_1, \varphi_2, \dots, \varphi_n$, $n = 2N + 1$, being the real trigonometric system $1, \cos 2\pi x, \sin 2\pi x, \dots, \cos N2\pi x, \sin N2\pi x$, we obtain that

$$\|\frac{1}{n}\sum_{j=1}^{n}\varphi_j - \sum_{j\in\Lambda}\frac{a_j}{m}\varphi_j\|_p \le C\gamma^{1/2}m^{-1/2}, \qquad \sum_{j\in\Lambda} a_j = m, \tag{2.10}$$

where $\sum_{j\in\Lambda}\frac{a_j}{m}\varphi_j$ is the $G^{i,\epsilon}_m(t)$. By (2.1) we find $\gamma \le p/2$. Next, by the Nikol'skii inequality we get from (2.10)

$$\|\frac{1}{n}\sum_{j=1}^n \varphi_j - \sum_{j\in\Lambda}\frac{a_j}{m}\varphi_j\|_\infty \le CN^{1/p}\|\frac{1}{n}\sum_{j=1}^n \varphi_j - \sum_{j\in\Lambda}\frac{a_j}{m}\varphi_j\|_p \le Cp^{1/2}N^{1/p}m^{-1/2}.$$

Choosing $p \asymp \ln N$ we obtain the desired in Theorem 1.2 bound.

We point out that the above proof of Theorem 1.2 gives the following statement.

Theorem 2.4. *Let $2 \le p < \infty$. For any $t \in \mathcal{RT}(N)$ the IA(ϵ) applied to $f := t/\|t\|_A$ provides after m iterations a m-term trigonometric polynomial $G_m(t) := G^{i,\epsilon}_m(f)\|t\|_A$ with the following approximation property*

$$\|t - G_m(t)\|_p \le Cm^{-1/2}p^{1/2}\|t\|_A, \quad \|G_m(t)\|_A = \|t\|_A,$$

with an absolute constant C.

3 Numerical integration and discrepancy

For a cubature formula $\Lambda_m(\cdot,\xi)$ we have

$$\Lambda_m(W^K_p,\xi) = \sup_{\|\varphi\|_{L_p(\Omega_y)}\le 1} |\int_{\Omega_y}\big(J(y) - \sum_{\mu=1}^m \lambda_\mu K(\xi^\mu,y)\big)\varphi(y)dy| =$$

$$= \|J(\cdot) - \sum_{\mu=1}^m \lambda_\mu K(\xi^\mu,\cdot)\|_{L_{p'}(\Omega_y)}. \tag{3.1}$$

Define the error of optimal cubature formula with m knots for a class W

$$\delta_m(W) := \inf_{\lambda_1,\dots,\lambda_m;\xi^1,\dots,\xi^m} \Lambda_m(W,\xi).$$

The above identity (3.1) obviously implies the following relation.

Proposition 3.1.

$$\delta_m(W^K_p) = \inf_{\lambda_1,\dots,\lambda_m;\xi^1,\dots,\xi^m} \|J(\cdot) - \sum_{\mu=1}^m \lambda_\mu K(\xi^\mu,\cdot)\|_{L_{p'}(\Omega_y)}.$$

Thus, the problem of finding the optimal error of a cubature formula with m knots for the class W^K_p is equivalent to the problem of best m-term approximation of a special function J with respect to the dictionary $\mathcal{D} = \{K(x,\cdot), x \in \Omega_x\}$.

Consider a problem of numerical integration of functions $K(x,y)$, $y \in \Omega_y$, with respect to x, $K \in \mathcal{K}_q$:

$$\int_{\Omega_x} K(x,y)dx - \sum_{\mu=1}^{m} \lambda_\mu K(\xi^\mu, y).$$

Definition 3.1. *(K,q)-discrepancy of a cubature formula Λ_m with knots $\xi^1,\dots,\xi^m$ and weights $\lambda_1,\dots,\lambda_\mu$ is*

$$D(\Lambda_m, K, q) := \| \int_{\Omega_x} K(x,y)dx - \sum_{\mu=1}^{m} \lambda_\mu K(\xi^\mu, y)\|_{L_q(\Omega_y)}.$$

The above definition of the (K,q)-discrepancy implies right a way the following relation.

Proposition 3.2.

$$\inf_{\lambda_1,\dots,\lambda_m;\xi^1,\dots,\xi^m} D(\Lambda_m, K, q)$$

$$= \inf_{\lambda_1,\dots,\lambda_m;\xi^1,\dots,\xi^m} \|J(\cdot) - \sum_{\mu=1}^{m} \lambda_\mu K(\xi^\mu, \cdot)\|_{L_q(\Omega_y)}.$$

Therefore, the problem of finding minimal (K,q)-discrepancy is equivalent to the problem of best m-term approximation of a special function J with respect to the dictionary $\mathcal{D} = \{K(x,\cdot), x \in \Omega_x\}$.

The particular case $K(x,y) = \chi_{[0,y]}(x) := \prod_{j=1}^{d} \chi_{[0,y_j]}(x_j), y_j \in [0,1), j = 1,\dots,d$, where $\chi_{[0,y]}(x)$, $y \in [0,1)$ is a characteristic function of an interval $[0,y)$, leads to a classical concept of the L_q-discrepancy.

Proof of Theorem 1.3. By (3.1)

$$\Lambda_m(W^K_p, \xi) = \|J(\cdot) - \sum_{\mu=1}^{m} \lambda_\mu K(\xi^\mu, \cdot)\|_{L_{p'}(\Omega_y)}.$$

We are going to apply Theorem 2.2 with $X = X(K,p') \subset L_{p'}(\Omega_y)$, $f = J_K$. We need to check the Condition B. Let F be a bounded linear functional on

$L_{p'}$. Then by the Riesz representation theorem there exists $h \in L_p$ such that for any $\phi \in L_{p'}$

$$F(\phi) = \int_{\Omega_y} h(y)\phi(y)dy.$$

By the Hölder inequality for any $x \in \Omega_x$ we have

$$\int_{\Omega_y} |h(y)K(x,y)|dy \le \|h\|_p.$$

Therefore, the functions $|h(y)K(x,y)|$ and $h(y)K(x,y)$ are integrable on $\Omega_x \times \Omega_y$ and by Fubini's theorem

$$F(J_K) = \int_{\Omega_y} h(y) \int_{\Omega_x} K(x,y)dx = \int_{\Omega_x} \left(\int_{\Omega_y} h(y)K(x,y)dy \right) dx$$

$$= \int_{\Omega_x} F(K(x,y))dx \le \sup_{x \in \Omega_x} F(K(x,y)),$$

which proves the Condition B. Applying Theorem 2.2 and taking into account (2.1) we complete the proof.

Proposition 3.2 and the above proof imply the following theorem on (K,q)-discrepancy.

Theorem 3.1. *Assume that* $K \in \mathcal{K}_q$ *satisfies the condition*

$$\|K(x,\cdot)\|_{L_q(\Omega_y)} \le 1, \quad x \in \Omega_x, \quad |\Omega_x| = 1$$

and $J_K \in X(K,q)$. *Then for any* m *there exists (provided by the Incremental Algorithm) a cubature formula* $\Lambda_m(\cdot,\xi)$ *with* $\lambda_\mu = 1/m$, $\mu = 1,2,\dots,m$, *and*

$$D(\Lambda_m, K, q) \le Cq^{1/2}m^{-1/2}, \quad 2 \le q < \infty.$$

References

[1] R.A. DeVore and V.N. Temlyakov, Nonlinear approximation by trigonometric sums, J. Fourier Analysis and Applications, **2** (1995), 29–48.

[2] Dilworth, S.J., Kutzarova, D., Temlyakov V.N., Convergence of some Greedy Algorithms in Banach spaces, *The J. Fourier Analysis and Applications* **8** (2002), 489–505.

[3] Donahue, M., Gurvits, L., Darken, C., Sontag, E., Rate of convex approximation in non-Hilbert spaces, *Constr. Approx.* **13** (1997), 187–220.

[4] Gluskin, E.D., Extremal properties of orthogonal parallelpipeds and their application to the geometry of Banach spaces, *Math USSR Sbornik* **64** (1989), 85–96.

[5] R.S. Ismagilov, Widths of sets in normed linear spaces and the approximation of functions by trigonometric polynomials, Uspekhi Mat. Nauk, **29** (1974), 161–178; English transl. in Russian Math. Surveys, **29** (1974).

[6] Maiorov, V.E., Trigonometric diameters of the Sobolev classes W_p^r in the space L_q, *Math. Notes* **40** (1986), 590–597.

[7] Temlyakov, V.N., Weak greedy algorithms, *Advances in Comput. Math.* **12** (2000), 213–227.

[8] Temlyakov, V.N., Greedy algorithms in Banach spaces, *Advances in Comput. Math.* **14** (2001), 277–292.

[9] Temlyakov, V.N., Cubature formulas and related questions, *J. Complexity* **19** (2003), 352–391.

[10] V.N. Temlyakov, Greedy-Type Approximation in Banach Spaces and Applications, Constr. Approx. **21** (2005), 257–292.

[11] V.N. Temlyakov, Greedy approximation, Cambridge University Press, 2011.