COUNTEREXAMPLES IN OPTIMAL QUADRATURE

Arthur G. Wershulz

CUCS-67-83

Division of Science and Mathematics
Fordham University/College at Lincoln Center
New York, N.Y. 10023

and

Department of Computer Science
Columbia University
New York, N.Y. 10027

July, 1983

This research was supported in part by the National Science Foundation under Grant MCS-8203271, and MCS-8303111.

Abstract

It is widely believed that order of exactness is a good measure of the quality of an algorithm for numerical quadrature. We show that this is not the case, by exhibiting a situation in which the optimal algorithm does not even integrate constants exactly. We also show that there are situations in which the penality for using equidistant nodes is unbounded. Finally, we show that the complexity of obtaining an ε-approximation can be an arbitrary function of ε, i.e., there is no hardest quadrature problem.

1. Introduction

In this Introduction, we use terms such as "algorithm," "information," "optimal," etc. without definition. They are rigorously defined in Section 2.

We consider the quadrature problem, and ask the following questions:

1. What is the relation between order of exactness and optimality of an algorithm?
2. Are equidistant or roughly-equidistant nodes nearly optimal?
3. If $\mathrm{comp}(\varepsilon)$ denotes the complexity of finding an ε-approximation, what kind of function can comp be?

We show the following:

1. There is no relation between order of exactness and optimality.
2. The penalty for using (roughly-) equidistant nodes can be unbounded.
3. The complexity, $\mathrm{comp}(\varepsilon)$, can be (almost) any non-decreasing function of ε such that $\lim_{\varepsilon \to 0} \mathrm{comp}(\varepsilon) = +\infty$.

In the remainder of this Introduction, we briefly discuss each of these issues in turn, citing the evidence that led us to ask these questions.

(1.) Order of exactness and optimality. Many people believe that maximizing the order of exactness of a quadrature rule (i.e., having it exactly integrate polynomials of as high a degree as possible) is one of the good ways to choose a quadrature rule, as may be seen by reading the chapter on quadrature in almost any numerical analysis textbook. (When the nodes are equally spaced, this leads to Newton-Cotes rules; when they are allowed to vary, this yields Gaussian rules.)

Furthermore, the literature abounds with situations in which some exactness condition is either imposed a priori or is necessary. As an example, suppose the integrand has $r \geq 1$ L_2-derivatives, the rth being given an a-priori L_2-bound. The Sard theory [8] only considers linear algorithms which are exact for polynomials of degree r - 1. However, this exactness condition need not be stated as an assumption. It can be derived from the facts that any homogeneous algorithm with finite error must integrate polynomials of degree r - 1 exactly, and that linear optimal algorithms exist, see [10, Chapter 3].

On the other hand, Motornyj [6] has shown that the rectangle rule (which is exact only for constants) is optimal for integrating periodic functions with $r \geq 1$ L_∞-derivatives. Hence, there is no general relation between the smoothness of

the integrand and the degree of exactness of the optimal rule. However, it may be shown that for this class of integrands, any homogeneous algorithm with finite error must integrate constants exactly.

Hence we see that in the examples presented above, any homogeneous algorithm must integrate constants exactly if it is to have finite error. Is this true in general?

In Section 3, we show that the answer to this problem is negative. In fact, we show an even stronger result. We construct classes of integrands for which the optimal algorithm does not integrate constants exactly, yet the nth optimal algorithm has finite error which is $\Theta(n^{-1})$ as $n \to \infty$.* Hence this algorithm provides a strong negative counterexample: it is not exact for constants, yet is optimal, has finite error, and is convergent.

Hence, there is no basic connection between order of exactness on the one hand, and finite error, convergence, or optimality on the other.

---

*In this paper, we use the classical O-notation, as well as Ω- and Θ-notation. If f and g are functions, then

$$f = \Omega(g) \quad \text{iff} \quad g = O(f)$$

and

$$f = \Theta(g) \quad \text{iff} \quad f = O(g) \text{ and } g = O(f).$$

(2.) Are roughly-equidistant nodes close to optimal? In all of the many classes of integrands considered in [10], roughly-equidistant nodes were (to within a constant) optimal points of evaluation for the integrand. Moreover, it is shown in [4] that equidistant nodes were optimal for the problem of integrating functions whose total variation was a priori bounded.

In this true in general, i.e., are roughly-equidistant nodes always almost-optimal? One reason to believe that this might be so is based on the fact that (as long as class of integrands is balanced and convex) an adaptive choice of nodes is never better than a nonadaptive choice, see [10,Chapter 2.6]. Hence, there appears to be no reason to do extra evaluations in some "critical" subinterval, since letting the points of evaluation depend on the integrand is no better than having them independent of the integrand. That is, it would be reasonable to expect that the points of evaluation should be more-or-less evenly distributed throughout the interval; that is, it should suffice to consider nodes which are roughly equidistant.

In Section 4, we show that this is not the case. We construct a class of integrands, as well as an optimal choice of evaluation points for that class which are not equidistant.

Moreover, we show that the penalty for using roughly-equidistant nodes is unbounded.

(3.) Is there a hardest quadrature problem? Bakhvalov has shown that for any $\alpha > 0$, there is a class of integrands for which the ε-complexity of quadrature is $\Theta(\varepsilon^{-\alpha})$ as $\varepsilon \to 0$. Is this the worst behavior for the ε-complexity? If not, is there a worst ε-complexity function, which would correspond to a hardest quadrature problem? In Section 5, we show that the answer is negative. In fact, the nth minimal error can go to zero arbitrarily slowly: for any "reasonable" function ψ for which $\psi(0) = 0$ and ψ is monotone increasing, there is a class of integrands for which the nth minimal error is $\Theta(\psi(\frac{1}{n}))$ as $n \to \infty$. This is used to show that the ε-complexity can be almost any nondecreasing function of ε such that $\lim_{\varepsilon \to 0} \text{comp}(\varepsilon) = +\infty$.

2. Preliminary concepts.

In this Section, we define the terminology used in the Introduction, and introduce some results from [10].

Let F be a balanced, convex subset of the Riemann-integrable functions on $[0,1]$. We wish to approximate the solution operator $S: F \to \mathbf{R}$ given by

$$(2.1) \qquad Sf := \int_0^1 f(t)dt \qquad \forall f \in F.$$

In what follows, we often refer to the problem (S,F). We suppose that all we know about an integrand is its values at a finite number of points in the interval. Hence, given a grid

$$(2.2) \qquad \Delta := \{0 \le t_1 < t_2 < \ldots < t_n \le 1\},$$

we only know the information

$$(2.3) \qquad N_\Delta f := [f(t_1) \ \ldots \ f(t_n)] \qquad \forall f \in F.$$

of cardinality n. An algorithm using this information is any mapping $\phi: N_\Delta(F) \to \mathbf{R}$; the class of all such algorithms is denoted by $\Phi(N_\Delta,F)$. The (worst-case) error of an algorithm ϕ is defined to be

$$(2.4) \qquad e(\phi,F) := \sup_{f\in F} |Sf - \phi(N_\Delta f)|.$$

We wish to find, for each information operator $N_\Delta: F \to \mathbf{R}^n$,

the best algorithm possible, i.e., the one with smallest error. Let

$$(2.5) \qquad e(N_\Delta, F) := \inf\{e(\varphi): \varphi \in \Phi(N_\Delta, F)\}$$

denote the optimal error from $\Phi(N_\Delta, F)$, i.e., the smallest error among all algorithms using N_Δ. Then results from [10] yield that

$$(2.6) \qquad e(N_\Delta, F) = \sup\{|Sz|: z \in F \cap \ker N_\Delta\}.$$

Moreover, Smolyak's lemma [3,10] tells us that there exist coefficients $\alpha_1, \ldots, \alpha_n \in \mathbf{R}$ such that

$$(2.7) \qquad e(\varphi^*, F) = e(N_\Delta, F),$$

where φ^* is the algorithm

$$(2.8) \qquad \varphi^*(N_\Delta f) := \sum_{i=1}^{n} \alpha_i f(t_i) \qquad \forall f \in F.$$

That is, there always exists a linear optimal error algorithm for F using N_Δ.

Next, we wish to choose, for any positive integer n, the best possible information of cardinality n. From (2.5), we see that this may be done as follows. Let

$$(2.9) \qquad e(n,F) := \inf_{\Delta} e(N_\Delta, F)$$

(the infinum being over all Δ satisfying (2.2)) be the n-th

minimal error. If there is a grid Δ^* satisfying (2.2) such that

$$(2.10) \qquad e(N_{\Delta^*}, F) = e(n, F),$$

then N_{Δ^*} is an n-th optimal information operator. Let $\mathfrak{A}_n(F)$ denote the class of all algorithms using information of cardinality n. If ϕ^* is an optimal error algorithm using the n-th optimal information N_{Δ^*}, then

$$(2.11) \qquad e(\phi^*, F) \leq e(\phi, F) \qquad \forall\, \phi \in \mathfrak{A}_n(F).$$

Hence, ϕ^* is an n-th minimal error algorithm.

Finally, we come to the notion of problem complexity, under the model of computation introduced in [10]. (The main features of this model are that each arithmetic operation has unit cost, and that evaluating any $f \in F$ at any point in $[0,1]$ has fixed finite cost c; we generally expect $c >> 1$ in practice.)

Let $\varepsilon > 0$ be a given error criterion. We define the problem complexity

$$(2.12) \qquad comp(\varepsilon, F) := \inf\{comp(\phi): e(\phi, F) \leq \varepsilon\},$$

$comp(\phi)$ denoting the complexity of the algorithm ϕ. That is, the problem complexity is the infinum of the complexities of all algorithms whose error does not exceed ε. Results from [10] yield that

(2.13) $\quad comp(\varepsilon,F) = (c+a)m(\varepsilon,F) - 1,$

where $a \in [1,2]$ and the ε-cardinality number $m(\varepsilon,F)$ is given by

(2.14) $\quad m(\varepsilon, F) := \inf\{n \in \mathbb{Z}: e(n,F) \leq \varepsilon\}.$

Roughly speaking, the ε-cardinality number is the functional inverse of the n-th minimal error. Since $c >> 1$, (2.13) and (2.14) imply (roughly speaking) that

(2.15) $\quad comp(\varepsilon,F) \approx ce^{-1}(\varepsilon).$

Moreover, the linear optimal error algorithm using n-th optimal information (where $n = m(\varepsilon,F)$) is (roughly speaking) an optimal complexity algorithm for solving the problem with error ε, since this algorithm has complexity at most $(c+2)m(\varepsilon,F) - 1$.

3. Exactness for constants is unnecessary.

In this Section, we show that a finite-error algorithm need not integrate constants exactly. This is done by showing an even stronger result: we construct a class of integrands for which the n-th minimal-error algorithm is not exact for constants, yet has error which is $\Theta(n^{-1})$ as $n \to \infty$.

To do this, we choose $r \in (0,1)$, define $\psi\colon [0,1] \to \mathbb{R}$ by

$$\psi(t) := t^r \qquad (0 \le t \le 1), \tag{3.1}$$

and let

$$G := \{g \in W^{1,\infty}(0,1)\colon g(0) = 0 \text{ and } \|g'\| \le 1\}, \tag{3.2}$$

where $W^{1,\infty}(0,1)$ is (as usual) the space of absolutely continuous g such that $g' \in L_\infty(0,1)$ and $\|\cdot\|$ is the $L_\infty(0,1)$-norm. Then we take

$$F := \{f:[0,1] \to \mathbb{R} \mid Rf \in G\} \tag{3.3}$$

as our class of integrands, where

$$(Rf)(x) := \frac{f(\psi^{-1}(x))}{\psi'(\psi^{-1}(x))} = f(\psi^{-1}(x))(\psi^{-1})'(x) \tag{3.4}$$

$$= \frac{1}{r} f(x^{1/r}) x^{1/r-1}.$$

Using the change of variables $t = \psi^{-1}(x)$, we have

$$(3.5) \qquad Sf = \int_0^1 f(t)dt = \int_0^1 (Rf)(x)dx = SRf \qquad \forall f \in F.$$

Hence, every element of F is Riemann-integrable. Morever, F contains nonzero constants.

As we will see below, the problem (S,G) is easy to analyze. However, we are more interested in the problem (S,F), since we will be able to construct optimal error algorithms for this problem which do not integrate constants exactly. Hence, we need to investigate the relations between these two problems.

We first consider information for the problem (S,F). Since there exist $f \in F$ such that $f(0)$ is infinite, we only consider information of the form

$$(3.6) \qquad N_\Delta f := [f(t_1) \ \ldots \ f(t_n)] \qquad \forall f \in F,$$

where

$$(3.7) \qquad \Delta := \{0 < t_1 < \ldots < t_n \leq 1\}.$$

For convenience, we let $t_0 := 0$ and $t_{n+1} := 1$; of course, $t_n = t_{n+1}$ when $t_n = 1$.

Next, we consider information for the problem (S,G). Since $g(0) = 0$ for any $g \in G$, we only consider information of the form

$$(3.8) \qquad N_\Delta g := [g(x_1) \ \ldots \ g(x_n)] \qquad \forall g \in G,$$

where

$$\tilde{\Delta} := \{0 < x_1 < \ldots < x_n \leq 1\}. \tag{3.9}$$

We let $x_0 := 0$ and $x_{n+1} := 1$.

The relation between the two problems is given in

Lemma 3.1: For information N_Δ for the problem (S,F) given by (3.6)-(3.7), let $N_{\tilde{\Delta}}$ be the information for the problem (S,G) given by (3.8)-(3.9) with

$$x_i := \psi(t_i) \qquad (1 \leq i \leq n). \tag{3.10}$$

(i) If

$$\varphi_{\tilde{\Delta}}(N_{\tilde{\Delta}} g) := \Sigma_{i=1}^{n} \alpha_i g(x_i) \qquad \forall g \in G \tag{3.11}$$

is a linear optimal error algorithm for G using $N_{\tilde{\Delta}}$, then

$$\varphi_\Delta(N_\Delta f) := \Sigma_{i=1}^{n} \frac{\alpha_i}{\psi'(t_i)} f(t_i) \qquad \forall f \in F \tag{3.12}$$

is a linear optimal error algorithm for F using N_Δ, and

$$e(\varphi_\Delta, F) = e(N_\Delta, F) = e(N_{\tilde{\Delta}}, G) = e(\varphi_{\tilde{\Delta}}, G). \tag{3.13}$$

(ii) The n-th minimal errors for the problems (S,F) and (S,G) are the same:

$$e(n,F) = e(n,G).$$

Moreover, N_Δ is n-th optimal information for (S,F) iff $N_{\tilde{\Delta}}$ is n-th optimal information for (S,G).

Proof: It suffices to prove (i), since (ii) follows immediately from (3.13). To show (i), let $f \in F$ and set $g := Rf$. Then (3.5) yields $Sf = Sg$, while (3.10)-(3.12) yield $\varphi_\Delta(N_\Delta f) = \varphi_{\tilde{\Delta}}(N_{\tilde{\Delta}} g)$. Hence $e(\varphi_\Delta, F) = e(\varphi_{\tilde{\Delta}}, G)$ by (2.4). Since $\varphi_{\tilde{\Delta}}$ is an optimal error algorithm for G using $N_{\tilde{\Delta}}$, $e(\varphi_{\tilde{\Delta}}, G) = e(N_{\tilde{\Delta}}, G)$. Finally, $e(N_{\tilde{\Delta}}, G) = e(N_\Delta, F)$ by (2.6) and (3.5). □

We are now able to exhibit a linear optimal error algorithm $\varphi_\Delta \in \Phi(N_\Delta, F)$. To do this, choose $\tilde{\Delta}$ as in Lemma 3.1 and define a space $L(\tilde{\Delta})$ of piecewise linear polynomials by

$$L(\tilde{\Delta}) := \{g \in W^{1,\infty}(0,1): g(0) = 0,\ g(x_n) = g(1),$$

$$g(x) = \xi_i x + \eta_i \text{ on } [x_i, x_{i+1}]\ (0 \le i \le n)\}. \tag{3.14}$$

Let $\{s_1, \ldots, s_n\}$ be the basis of $L(\tilde{\Delta})$ for which

$$s_i(x_j) = \delta_{ij} \qquad (0 \le j \le n+1,\ 1 \le i \le n+1)$$

$$s_n(x_j) = 0 \qquad (0 \le j \le n-1) \tag{3.15}$$

$$s_n(x_n) = s_n(1) = 1$$

and set

$$(3.16) \qquad \beta_i := \frac{1}{r}\, t_i^{1-r} \int_0^1 s_i(x)\,dx \qquad (1 \le i \le n).$$

We then define

$$(3.17) \qquad \varphi_\Delta(N_\Delta f) := \Sigma_{i=1}^{n} \beta_i f(t_i).$$

Theorem 3.1:

(i) φ_Δ is a linear optimal error algorithm for F using N_Δ.

(ii) Let

$$\delta_i := x_{i+1} - x_i \qquad (0 \le i \le n).$$

Then

$$(3.18) \qquad e(\varphi_\Delta, F) = e(N_\Delta, F) = e(N_{\widetilde{\Delta}}, G) = \frac{1}{4} \Sigma_{i=0}^{n-1} \delta_i^2 + \frac{1}{2} \delta_n^2.$$

Proof: (i) For $g \in G$, we define the interpolant $P_{\widetilde{\Delta}} g \in L(\widetilde{\Delta})$ by

$$(3.19) \qquad (P_{\widetilde{\Delta}} g)(x) := \Sigma_{i=1}^{n} g(x_i) s_i(x).$$

Then (3.11) and (3.16) yield

$$(3.20) \qquad \varphi_{\widetilde{\Delta}}(N_{\widetilde{\Delta}} g) = S P_{\widetilde{\Delta}} g.$$

Following the techniques of [9, pp. 30-31], we see that $\varphi_{\widetilde{\Delta}}$ is a linear optimal error algorithm for G using $N_{\widetilde{\Delta}}$. Hence, Lemma 3.1 yields that φ_Δ is a linear optimal error algorithm for F using N_Δ.

(ii) It suffices to prove the last equality. Define

$$(3.21)\qquad z_{\underset{\sim}{\Delta}}(x) := \begin{cases} x - x_i & \text{for } x \in [x_i, \frac{1}{2}(x_i+x_{i+1})], \quad 0 \le i \le n \\ x_{i+1}-x & \text{for } x \in [\frac{1}{2}(x_i+x_{i+1}), x_{i+1}], \quad 0 \le i \le n \\ x-x_n & \text{for } x \in [x_n, 1]. \end{cases}$$

As in [9, pp. 55-57, we see that

$$(3.22)\qquad e(N_{\underset{\sim}{\Delta}}, G) = Sz_{\underset{\sim}{\Delta}}.$$

A short calculation yields

$$(3.23)\qquad Sz_{\underset{\sim}{\Delta}} = \frac{1}{4} \sum_{i=0}^{n-1} \delta_i^2 + \frac{1}{2}\delta_n^2.$$

The result follows from (3.22) and (3.23). □

Remark 3.1: The proof of Theorem 3.1 establishes an even stronger result. An algorithm is a strongly optimal error algorithm if it produces a best possible approximation to Sf for every $f \in F$ (rather than for just a worst-case $f \in F$). Since $\varphi_{\underset{\sim}{\Delta}}$ is a strongly optimal error algorithm for G using $N_{\underset{\sim}{\Delta}}$, we find that φ_Δ is a strongly optimal error algorithm for F using N_Δ. □

We now show that for $r \in (\frac{1}{2}, 1)$, the optimal error algorithm $\varphi_\Delta \in \Phi(N_\Delta, F)$ has no "order of exactness," i.e., it does not even integrate (nonzero) constants exactly.

Theorem 3.2: Suppose that $r \in (\frac{1}{2}, 1)$. Let $|c| \le \frac{r^2}{|1-r|}$, so that the constant function $f: [0,1] \to \mathbb{R}$ given by

$$f(t) := c \qquad (0 \le t \le 1)$$

belongs to F. Then

$$\operatorname{sgn}(Sf - \varphi_\Delta(N_\Delta f)) = \operatorname{sgn} c.$$

Hence, φ_Δ does not integrate nonzero contstants exactly.

Proof: If $c = 0$, then linearity of φ_Δ yields

$$(3.24) \qquad \varphi_\Delta(N_\Delta f) = 0 = Sf.$$

We now consider the case $c > 0$. Let $g = Rf$, i.e.,

$$(3.25) \qquad g(x) = \frac{c}{r}\, x^{1/r-1}.$$

Then $r \in (\frac{1}{2}, 1)$ implies that g is monotone increasing and strictly concave on $[0,1]$. Hence

$$(3.26) \qquad g(x) > (P_\Delta g)(x) \qquad \text{a.e.} \quad x \in [0,1].$$

So (3.5), (3.10)-(3.12), and (3.20) yield

$$(3.27) \qquad Sf - \varphi_\Delta(N_\Delta f) = Sg - SP_\Delta g > 0,$$

as required.

Finally, when $c < 0$, the same technique yields

$$(3.28) \qquad Sf < \varphi_\Delta(N_\Delta f). \qquad \square$$

We now determine nth optimal information for F. Let

Δ_n^* and $\tilde{\Delta}_n^*$ denote the grids

$$(3.29) \qquad \Delta_n^* := \{0 < t_1^* < \ldots < t_n^* < 1\} \text{ with } t_i^* := \left(\frac{2i}{2n+1}\right)^{1/r} \qquad (1 \le i \le n)$$

and

$$(3.30) \qquad \tilde{\Delta}^*_n := \{0 < x_1^* < \ldots < x_n^* < 1\} \text{ with } x_i^* := \frac{2i}{2n+1} \qquad (1 \le i \le n),$$

respectively.

Theorem 3.3:

$$e(n,F) = e(n,G) = \frac{1}{4n+2} .$$

Moreover,

$$e(N_{\tilde{\Delta}},G) = e(n,G) \quad \text{iff} \quad \tilde{\Delta} = \tilde{\Delta}_n^*,$$

and

$$e(N_{\Delta},F) = e(n,F) \quad \text{iff} \quad \Delta = \Delta_n^*.$$

Proof: Immediate from Lemma 3.1, Theorem 3.1 and minimizing the right-hand side of (3.18) subject to the constraint $\Sigma_{i=0}^{n} s_i = 1$. □

Hence, sampling at the points $\{t_1^*,\ldots,t_n^*\}$ is nth optimal information for the class F of integrands. Note that the nth minimal error is finite for all $n \ge 0$, and is $\Theta(n^{-1})$ as $n \to \infty$. Thus we have found a situation in which the nth minimal

error algorithms all have finite error and are convergent, but do not integrate constants exactly.

Remark 3.2: We have shown that ω_Δ does not integrate constants exactly when $r \in (\frac{1}{2},1)$. One can show that the nth minimal error algorithm $\omega_{\Delta_n^*}$ does not integrate constants exactly when $r \in [\frac{1}{4},\frac{1}{3}] \cup [\frac{1}{2},1]$. We do not know whether this assumption on r is needed. In fact, numerical testing indicates that the assumption may not be necessary. □

We now discuss some open questions.

Problem 3.1: Although we know that the nth minimal-error algorithm does not integrate constants exactly, we do not know what the penalty is for using a quadrature rule that integrates constants exactly. For instance, if $\{\omega_n\}_{n=1}^\infty$ is a sequence of quadrature rules which integrate constants exactly, and there exists $M > 0$ such that $e(\varphi_n,F) \le Mn^{-1}$ for all sufficiently large n, then the complexity penalty for using $\{\omega_n\}_{n=1}^\infty$, instead of $\{\omega_{\Delta_n^*}\}_{n=1}^\infty$, in computing an ε-approximation is bounded as $\varepsilon \to 0$. In such a case, it may be worthwhile to use the (perhaps simpler) rule ω_n, rather than $\omega_{\Delta_n^*}$. Does such a sequence $\{\omega_n\}_{n=1}^\infty$ exist? □

Problem 3.2: Although $\omega_{\Delta_n^*}$ does not integrate constants exactly, ω_Δ integrates the functions $f(t) = t^{r-1}$ and $f(t) = t^{2r-1}$

exactly when $t_n = 1$. Is there a function f which is integrated exactly by $\omega_{\Delta_n^*}$ for all n? If so, does there exist a different class F of integrands for which the linear optimal error algorithm ω_n using nth optimal information is convergent, yet for which the only element of F which is exactly integrated by ω_n for all (sufficiently large n) is the zero function? □

4. Roughly-equidistant nodes can be terrible.

In this Section, we show that the penalty for sampling at almost equidistant toints (rather than optimal points) is unbounded, whether the penalty is measured by error or complexity.

We define F as in Section 3, but now with $0 < r < \frac{1}{2}$. Recall that Theorem 3.3 states that the nth minimal error is $\Theta(n^{-1})$ as $n \to \infty$, and that sampling at the points $\{t_1^*,\ldots,t_n^*\}$ with $t_i^* = (\frac{2i}{2n+1})^{1/r}$ $(1 \le i \le n)$ is optimal.

We now ask how close sampling at "roughly-equidistant" nodes is to optimal. In crder to make this precise, we will use information based on a sequence of grids $\{\Delta_n\}_{n=1}^{\infty}$ which is <u>quasi-uniform</u> (see e.g. [7].) That is,

$$\limsup_{n\to\infty} \mu(\Delta_n) < +\infty \tag{4.1}$$

where

$$\Delta_n := \{0 < t_{1,n} <\ldots< t_{n,n} \le 1\}$$

$$t_{0,n} := 0 \qquad t_{n+1,n} := 1$$

$$h_{i,n} := t_{i+1,n} - t_{i,n}$$

$$h(\Delta_n) := \max_{0\le i\le n} h_{i,n} \tag{4.2}$$

$$\mu(\Delta_n) := \begin{cases} \max\limits_{0\le i\le n} \dfrac{h(\Delta_n)}{h_{i,n}} & \text{if } t_{n,n} \ne 1 \\ \max\limits_{0\le i\le n-1} \dfrac{h(\Delta_n)}{h_{i,n}} & \text{if } t_{n,n} = 1 \end{cases}$$

(For example, a sequence of equidistant grids is quasi-uniform.)

Remark 4.1: The sequence $\{\Delta_n^*\}_{n=1}^{\infty}$ of optimal grids is not quasi-uniform. However, Δ_n^* is the image of $\tilde{\Delta}_n^*$ under the transformation ψ, and $\{\tilde{\Delta}_n^*\}_{n=1}^{\infty}$ is quasi-uniform. That is, even though the optimal grid sequence is not quasi-uniform, there is a change of variables $x = t^r$ such that the transformed grid sequence is quasi-uniform. Of course, this is precisely what is done in practice to remove the singularity at the origin. □

We then have

Theorem 4.1: $e(N_{\Delta_n}, F) = \Theta(n^{-2r})$.

Proof: Let

$$\tilde{\Delta}_n := \{0 < x_{1,n} < \ldots < x_{n,n} \leq 1\} \tag{4.3}$$

with

$$x_{i,n} := \psi(t_{i,r}) = t_{i,n}^r \quad (0 \leq i \leq n+1), \tag{4.4}$$

and let

$$\xi_{i,n} := x_{i+1,n} - x_{i,n} \quad (0 \leq i \leq n). \tag{4.5}$$

Then Theorem 3.2 yields

$$(4.7) \qquad e(N_{\Delta_n}, F) = \Theta(\Sigma_{i=0}^{n} \delta_{i,n}^2).$$

In what follows, we assume that $\delta_{n,n} > 0$; the grids for which $\delta_{n,n} = 0$ are handed analogously.

We next estimate $\delta_{0,n}, \ldots, \delta_{n,n}$. By quasi-uniformity of $\{\Delta_n\}_{n=1}^{\infty}$, we find that there exists $\alpha \geq 1$ such that

$$(4.8) \qquad \frac{1}{\alpha}\frac{1}{n+1} \leq h_{i,n} \leq \frac{\alpha}{n+1}$$

and so

$$(4.9) \qquad \frac{i}{\alpha(n+1)} \leq t_{i,n} \leq \frac{i\alpha}{n+1}.$$

We first consider $\delta_{0,n} = h_{0,n}^r$. Then

$$(4.10) \qquad \frac{1}{\alpha^r (n+1)^r} \leq \delta_{0,n} \leq \frac{\alpha^r}{(n+1)^r}.$$

Now let $i \in \{1, \ldots, n\}$. By the mean-value theorem

$$(4.11) \qquad \delta_{i,n} = \psi'(\tau_{i,n}) h_{i,n},$$

where

$$(4.12) \qquad \tau_{i,n} \in (t_{i,n}, t_{i+1,n}) \subseteq \left[\frac{i}{\alpha(n+1)}, \frac{\alpha(i+1)}{n+1}\right].$$

Since $\psi'' < 0$ on $(0,1)$, (4.11) and (4.12) yield

$$(4.13) \qquad \psi'\left(\frac{\alpha(i+1)}{n+1}\right) h_{i,n} \leq \delta_{i,n} \leq \psi'\left(\frac{i}{\alpha(n+1)}\right) h_{i,n}.$$

Hence (3.1), (4.8) and (4.13) yield

$$(4.14)\qquad \frac{r}{\alpha^{2-r}}\left(\frac{1}{n+1}\right)^{r}\left(\frac{1}{i+1}\right)^{1-r} \le \delta_{i,n} \le r\alpha^{2-r}\left(\frac{1}{n+1}\right)^{r}\left(\frac{1}{i}\right)^{1-r} \qquad (1 \le i \le n).$$

Let

$$\sigma_{n,r} = \Sigma_{i=1}^{n}\left(\frac{1}{i}\right)^{2-2r}$$

$$(4.15)\qquad C_1(\alpha,r) := \frac{r^2}{\alpha^{4-2r}}$$

$$C_2(\alpha,r) := \max\{r^2\alpha^{4-2r}, \alpha^{2r}\}.$$

Then (4.8), (4.14) and (4.15) yield

$$(4.16)\qquad \frac{C_1(\alpha,r)}{(n+1)^{r}}(1+\sigma_{n,r}) \le \Sigma_{i=1}^{n}\delta_{i,n}^{2} \le \frac{C_2(\alpha,r)}{(n+1)^{r}}(1+\sigma_{n,r}).$$

Since $r < \frac{1}{2}$, $\sigma_{n,r}$ is a convergent series of positive terms, i.e., there exists $\sigma_r > 0$ for which $\sigma_{n,r} \le \sigma_r$ for all n. Hence (4.16) becomes

$$(4.17)\qquad \frac{C_1(\alpha,r)}{(n+1)^{r}} \le \Sigma_{i=0}^{n}\delta_{i,n}^{2} \le \frac{C_2(\alpha,r)}{(n+1)^{r}}(1+\sigma_r).$$

The result now follows from (4.7) and (4.17). □

As an immediate consequence, we see the penalty, in terms of optimal error, for using quasi-uniform nodes:

Corollary 4.1:

$$\frac{e(N_n,F)}{e(n,F)} = \Theta(n^{1-2r}) \to +\infty \qquad \text{as} \quad n \to \infty. \quad \square$$

We next investigate the penalty, in terms of complexity, for using quasi-uniform information. Let $comp^*(\varepsilon,F)$ denote the ε-complexity when using information based on a quasi-uniform grid sequence, i.e.,

$$(4.18) \qquad comp^*(\varepsilon,F) := \inf\{comp(\omega): e(\omega,F) \le \varepsilon \text{ and } \omega \in \bigcup_{n=1}^{\infty} \Phi(N_{\Delta_n},F)\}.$$

Then results from [10] yield that

$$(4.19) \qquad comp^*(\varepsilon,F) = \Theta(m^*(\varepsilon,F)) \qquad \text{as} \quad \varepsilon \to 0,$$

where

$$(4.20) \qquad m^*(\varepsilon,F) := \inf\{n: e(N_{\Delta_n},F) \le \varepsilon\}$$

may be thought of as an ε-cardinality number for information based on a quasi-uniform grid sequence.

We then have

Corollary 4.2:

$$\frac{comp^*(\varepsilon,F)}{comp(\varepsilon,F)} = \Theta\left(\left(\frac{1}{\varepsilon}\right)^{1/(2r)-1}\right) \to \infty \qquad \text{as} \quad \varepsilon \to 0.$$

Proof: From (4.20) and Theorem 4.1, we have

$$(4.21) \qquad m^*(\varepsilon,F) = \Theta\left(\left(\frac{1}{\varepsilon}\right)^{1/(2r)}\right) \qquad \text{as} \quad \varepsilon \to \infty.$$

Hence (4.19) and (4.21) yield

$$(4.22) \qquad \mathrm{comp}^*(\epsilon,F) = \Theta((\frac{1}{\epsilon})^{1/2r}) \qquad \text{as} \quad \epsilon \to \infty.$$

On the other hand, Theorem 3.3 implies that

$$(4.23) \qquad m(\epsilon,F) = \Theta(\frac{1}{\epsilon}) \quad \text{as} \quad \epsilon \to \infty.$$

Hence (2.13) and (4.23) imply

$$(4.24) \qquad \mathrm{comp}(\epsilon,F) = \Theta(\frac{1}{\epsilon}) \qquad \text{as} \quad \epsilon \to \infty.$$

The Corollary follows from (4.22) and (4.24). □

Hence, Corollaries 4.1 and 4.2 tell us that the penalty for using information based on a quasi-uniform sequence of grids, whether measured by error or complexity, is unbounded. However, the optimal grid sequence becomes quasi-uniform after a change of variables. This leads to an open question:

Problem 4.1: Does there exist a bounded convex class F of integrands such that the nth minimal error is finite for all n and converges to zero as $n \to \infty$, and for which there is no change of variables under which the sequence of optimal grids becomes quasi-uniform? □

5. Quadrature is Arbitrarily Hard

In this Secion we show that F may be chosen so that the nth minimal error goes to zero as slowly as we like. Hence, the ε-complexity can be a function which goes to infinity (for $\varepsilon \to 0$) as quickly as we desire. See [1] for a related result.

To this end, we let ψ: $[0,2] \to \mathbf{R}$ satisfy

$$\psi \in C[0,2] \cap C^1(0,2],$$

$$\psi' > 0 \quad \text{on} \quad (0,2],$$

(5.1)

$$\psi(0) = 0, \quad \text{and}$$

$$\limsup_{t\to\infty} \frac{\psi(t)}{t} > 0 \qquad \text{(possibly infinite).}$$

Define a seminorm $|\cdot|_\psi$ and a norm $\|\cdot\|_\psi$ on $C^1[0,1]$ by

$$|f|_\psi := \operatorname*{ess\,sup}_{0\le t<t'\le 1} \frac{|f(t')-f(t)|}{\psi(2t-2t')} \tag{5.2}$$

and

$$\|f\|_\psi := \max\{\|f\|_{L_\infty(0,1)}, |f|_\psi\}. \tag{5.3}$$

Let W_ψ be the Banach space given by the $\|\cdot\|_\psi$-closure of $C^1[0,1]$ in $L_\infty(0,1)$. Finally, let

$$F := \{f \in W_\psi : \|f\|_\psi \le 1\}. \tag{5.4}$$

Remark 5.1: Our motivating example is given by $\psi(t) := t^{\alpha}$ $(0 < \alpha < 1)$, in which case $W_{\psi} = \text{Lip } \alpha$, (see e.g. [5]). Another ψ satisfying (5.1) is $\psi(t) := (\ell n \frac{1}{t})^{-\alpha}$, where $\alpha > 0$. Note that (5.1) tells us that ψ is a modulus of continuity, see [6]. □

Our results will be first stated in terms of the function $\rho: [0,1] \to \mathbb{R}$ given by

$$\rho(h) := \begin{cases} \frac{1}{h} \int_0^h \psi(t)dt & h > 0 \\ 0 & h = 0 \end{cases} \tag{5.5}$$

By L'Hospital's rule, ρ is continuous. We will first show that $e(n,F) = \rho(\frac{1}{n})$. We will then show a mild smoothness condition on ψ which will imply that $\rho(h) = \Theta(\psi(h))$ as $h \to 0$. This will yield the result $e(n,F) = \Theta(\psi(\frac{1}{n}))$ as $n \to \infty$.

Remark 5.2: We briefly comment on (5.1). Recall that $\psi(\frac{1}{n})$ will (essentially) be the nth minimal error. The first three conditions are equivalent to the requirement that the nth minimal error decrease as n increases, going to zero as $n \to \infty$. The final condition tells us that the nth minimal error is $\Omega(n^{-1})$ as $n \to \infty$ (i.e. the problems are not too easy); moreover, if it does not hold, then W_{ψ} consists only of constant functions. □

We first determine the optimal error, the nth minimal error, the nth optimal information, and the nth minimal error algorithm in

Theorem 5.1:

(i) The optimal error among algorithms using N_Δ is

$$e(N_\Delta, F) = \frac{1}{2} \int_0^{2h_0} \psi + \Sigma_{i=1}^{n-1} \int_0^{h_i} \psi + \frac{1}{2} \int_0^{2h_n} \psi, \tag{5.6}$$

where

$$h_i := t_{i+1} - t_i \qquad (0 \le i \le n)$$

with

$$t_0 := 0 \quad \text{and} \quad t_{n+1} := 1.$$

(ii) The nth minimal error is

$$e(n, F) = \rho(\frac{1}{n}) \tag{5.7}$$

and

$$N_n^* f := [f(t_1^*) \ \ldots \ f(t_n^*)), \tag{5.8}$$

with

$$t_i^* = \frac{2i-1}{2n} \qquad (1 \le i \le n) \tag{5.9}$$

is the unique n-th optimal information.

(iii) The composite midpoint rule

$$\varphi_n(N_n^* f) := \frac{1}{n} \Sigma_{i=1}^{n} f(t_i^*) \tag{5.10}$$

(with t_i as in (5.9)) is an nth minimal error algorithm, with

$$(5.11) \qquad e(\varphi_n, F) = \rho(\tfrac{1}{n}).$$

Proof (i): Let

$$(5.12)\quad z_\Delta(t) := \begin{cases} \psi(2t_1-2t) & 0 \le t \le t_1 \\ \psi(2t-2t_1) & t_1 \le t \le t_1 + \frac{1}{2}h_1 \\ \psi(2t_i-2t) & t_i - \frac{1}{2}h_{i-1} \le t \le t_i \\ \psi(2t-2t_i) & t_i \le t \le t_i + \frac{1}{2}h_i \\ \psi(2t_n-2t) & t_n - \frac{1}{2}h_{n-1} \le t \le t_n \\ \psi(2t-2t_n) & t_n \le t \le 1 \end{cases} \quad (2 \le i \le n-1)$$

Then $z_\Delta \in F \cap \ker N_\Delta$ and

$$(5.13) \qquad -z_\Delta(t) \le z(t) \le z_\Delta(t) \quad (0 \le t \le 1) \ \forall\, z \in F \cap \ker N_\Delta.$$

As in the proof of Theorem 3.1 (5.13) yields,

$$(5.14) \qquad e(N_\Delta, F) = \int_0^1 z_\Delta(t)\,dt .$$

So (5.6) follows from (5.12) and (5.14) by a straightforward calculation.

(ii) follows by minimizing the right-hand side of (5.6) subject to the constraint $\Sigma_{i=0}^{n} h_i = 1$.

(iii) Let $f \in F$ and set $h = \frac{1}{2n}$. Then

$$(5.15)\qquad \int_{t_i^*-h}^{t_i^*} |f(t)-f(t_i^*)|\,dt \le \int_{t_i^*-h}^{t_i^*} \psi(2t_i^*-2t)\,dt = \frac{1}{2}\int_0^{1/n} \psi(t)\,dt$$

and, analogously,

$$(5.16)\qquad \int_{t_i^*}^{t_i^*+h} |f(t)-f(t_i^*)|\,dt \le \frac{1}{2}\int_0^{1/n} \psi(t)\,dt.$$

So (5.5), (5.15) and (5.16) yield

$$(5.17)\qquad \left|\int_0^1 f(t)\,dt - \varphi_n(N_n^* f)\right| \le 2n\cdot\frac{1}{2}\int_0^{1/n} \psi(t)\,dt = n\int_0^{1/n} \psi(t)\,dt = \rho\left(\frac{1}{n}\right).$$

Since $f \in F$ is arbitrary, we find $e(\varphi_n,F) \le \rho(\frac{1}{n})$. On the other hand, $\varphi_n \in \Phi(N_n^*,F)$, so that $e(\varphi_n,F) \ge e(N_n^*,F) = \rho(\frac{1}{n})$. Hence, $e(\varphi_n,F) = \rho(\frac{1}{n})$, completing the proof. □

We now turn to complexity. Using (2.13) and (2.14), the results of Theorem 5.1 immediately yield

Theorem 5.2:

(i) The ε-complexity satisfies

$$(5.18)\qquad (c+1)m(\varepsilon,F) - 1 \le \mathrm{comp}(\varepsilon,F) \le (c+1)m(\varepsilon,F)$$

where

$$(5.19) \qquad m(\varepsilon,F) = \left\lceil \frac{1}{\left|\rho^{-1}(\varepsilon)\right|} \right\rceil .$$

(ii) φ_n, with $n = m(\varepsilon,F)$, is an algorithm for which

$$(5.20) \qquad e(\varphi_n,F) \le \varepsilon \quad \text{and} \quad \text{comp}(\varphi_n) \le (c+1)\, m(\varepsilon,F). \qquad \square$$

Hence φ_n with $n = m(\varepsilon,F)$ is (disregarding one arithmetic operation) an optimal complexity algorithm for ε-approximation, and

$$(5.21) \qquad \text{comp}(\varepsilon,F) \sim \frac{c+1}{\left|\rho^{-1}(\varepsilon)\right|} \qquad \text{as} \quad \varepsilon \to 0.$$

Finally, we wish to translate the dependence in Theorems 5.1 and 5.2 on ρ into dependence on the original function ψ. To do this, we impose a mild smoothness condition on ψ, namely, that

$$(5.22) \qquad \beta := \lim_{h\to 0} \frac{h\psi'(h)}{\psi(h)}$$

exists and is finite. (Note that this holds for the functions mentioned in Remark 5.1.) Then L'Hospital's rule gives

$$(5.23) \qquad \lim_{h\to\infty} \frac{\rho(h)}{\psi(h)} = \frac{1}{1+\beta}.$$

Using (5.23) and Theorems 5.1 and 5.2, we have

Theorem 5.3:

(i) $e(n,F) \sim \frac{1}{1+\beta}\, \psi(\frac{1}{n}) \qquad \text{as} \qquad n \to \infty.$

(ii) $\mathrm{comp}(\varepsilon, F) \sim \dfrac{c+1}{\psi^{-1}((1+\beta)\varepsilon)}$ as $\varepsilon \to 0$. □

Since ψ is essentially an arbitrary function, this tells us that quadrature can be arbitrarily hard.

Acknowledgements

I wish to thank D. Lee and J.F. Traub (Columbia University), as well as G. Wasilkowski (University of Warsaw and Columbia University) for their most valuable comments and suggestions.

References

[1] Babenko, V.F., "Asymptotically sharp bounds for the remainder for the best quadrature formulas for several classes of functions," Math. Notes 19 (1976), 187-193.

[2] Bakhvalov, N.S., "On the approximate calculation of multiple integrals" (in Russian), Vestnik MGU Ser. Math. Mech. Astron. Phys. Chem. 4 (1959), 3-18.

[3] Bakhvalov, N.S., "On the optimality of linear methods for operator approximation in convex classes of functions," USSR Computational Math. and Math. Phys. 11 (1971), 244-249.

[4] Lee, D. and Traub, J.F., "Optimal integration for functions of bounded variation," Research Report, Computer Science Department, Columbia University, New York, 1983.

[5] Lorentz, G.G., Approximation of Functions. Holt, Rinehard, and Winston, New York, 1966.

[6] Motornyj, V.P., "On the best quadrature formulas of the form $\Sigma_{n=1}^{n} p_k f(x_k)$ for some classes of periodic differentiable functions" (in Russian), Dokl. Akad. Nauk SSSR Ser. Math. 38 (1974), 583-614.

[7] Oden, J.T. and Reddy, J.N., An Introduction to the Mathematical Theory of Finite Elements, Wiley-Interscience, New York, 1976.

[8] Sard, A., Linear Approximation. American Mathematical Society, Providence, Rhode Island, 1963.

[9] Traub, J.F., Wasilkowski, G., and Woźniakowski, H., Information, Uncertainty, Complexity. Addison-Wesley, Reading, Massachusetts, 1983.

[10] Traub, J.F. and Woźniakowski, H., A General Theory of Optimal Algorithms. Academic Press, New York, 1980.