STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume XLIV, Number 1, March 1999

ALMOST OPTIMAL NUMERICAL METHOD

SOMOGYI ILDIKO

Abstract. This paper investigates an algorithm presented by Smolyak (1963), who studied tensor product problems.

1. Introduction

The essence of these algorithms is that it is enough to know how to solve the tensor product problem for $d = 1$ efficiently. The algorithms for arbitrary d are fully determined in terms of the algorithms for generally, arbitrary linear functionals.

The choice of function values is especially interesting, since for arbitrary linear functionals we know how to solve multivariate problems.

The algorithms are linear. They depend linearly on the information. This property makes their implementation easier. In fact, the weights of the algorithm for $d \geq 2$ are given by linear combinations of the corresponding tensor product weights of the one dimensional algorithms. Information used by the algorithms is called hyperbolic cross information and had been successfully applied for a number of problems.

2. Formulation of the problem

In this section a tensor product problem will be define for a class of functions of d variables.

For $d = 1, 2, \ldots$ consider

$$S_d : X_d \to Y_d$$

where X_d is a separable Banach space of functions $f : D^d \to \mathbf{R}$, $D \subset \mathbf{R}$, Y_d is either a separable Hilbert space of functions, or $\mathbf{R}$, and S_d is a continuous linear operator.

1991 *Mathematics Subject Classification:* 65J99.

Key words and phrases: tensor product, numerical methods, optimality.

We assume that Y_d is a tensor product,

$$Y_d = Y_1 \otimes Y_1 \otimes \ldots \otimes Y_1, \tag{1}$$

and X_1 is a Hilbert space

$$X_d = X_1 \otimes X_1 \otimes \ldots \otimes X_1$$

$$S_d = S_1 \otimes S_1 \otimes \ldots \otimes S_1.$$

The tensor product $f = f_1 \otimes \ldots \otimes f_d = \bigotimes_{k=1}^{d} f_k$ for numbers f_k is just the product $\prod_{k=1}^{d} f_k$. When f_k are scalar functions, f is a function of d variables, $f(t_1, \ldots, t_d) = \prod_{k=1}^{d} f_k(t_k)$.

The element $S_d(f)$ is approximated by $A(f) = \phi(N(f))$, where the information about f,

$$N(f) = [L_1(f), \ldots, L_n(f)], \tag{2}$$

consists of n values of continuous linear functionals L_i, and $\phi : \mathbf{R}^n \to G_d$ is a linear mapping. This results from linearity of A,

$$A(f) = \sum_{i=1}^{n} y_i L_i(f), \qquad \text{for some } y_i \in Y_d. \tag{3}$$

The error of the algorithm A is given as

$$e(A) = \sup \{ ||S_d(f) - A(f)||_{Y_d} \ : \ ||f||_{X_d} \leq 1 \}. \tag{4}$$

Due to linearity of S_d and A, we have

$$e(A) = ||S_d - A||.$$

The cost of A does not depend on the setting and it is defined as follows. We assume that the cost of computing $L_i(f)$ equals $c(d)$ for any $f \in X_d$ and any L_i. Also assume that basic arithmetic operations on reals and multiplication and addition in Y_d have a unit cost. Assuming that the elements y_i can be precomputed, the cost of the algorithm A, $cost(A)$, is bounded by

$$cost(A) \leq n(c(d) + 2) - 1.$$

The precomputation of the elements y_i is usually easy since they depend only on the corresponding elements for $d = 1$.

3. Smolyak's algorithm

As it was mentioned in the introduction, the essence of these algorithms is that they give a general construction that leads to almost optimal approximations for any dimension $d > 1$ from optimal approximation for the univariate case $d = 1$.

Assume, therefore, that for $d = 1$, we know linear algorithms (operators) U^i, $i \geq 1$, which approximate the problem $\{X_1, Y_1, S_1\}$ such that $||S_1 - U^i|| \to 0$ as $i \to \infty$. Introducing the notation

$$\Delta_0 = U_0 = 0, \qquad \Delta_i = U_i - U_{i-1}, \tag{5}$$

for $d > 1$ we approximate the tensor product problem $\{X_d, Y_d, S_d\}$ by the algorithm

$$A(q,d) = \sum_{0 \leq i_1 + i_2 + \cdots + i_d \leq d} \Delta_{i_1} \otimes \ldots \otimes \Delta_{i_d}. \tag{6}$$

Hence $f(t_1, t_2, ...t_d) = f_1(t_1)f_2(t_2)\ldots f_d(t_d)$ then

$$(A(q,d)f)(t_1, t_2, \ldots t_d) = \sum_{0 \leq i_1 + i_2 + \cdots + i_d \leq d} (\Delta_{i_1} f_1)(t_1)(\Delta_{i_2} f_2)(t_2)\ldots(\Delta_{i_d} f_d)(t_d)$$

where q is a nonnegativ integer, and $q \geq d$, because when $q < d$ one of the indices is zero, say $i_j = 0$, and $\Delta_{i_j} = 0$ implies that $A(q,d) = 0$.

We use the notation $|i| = i_1 + \cdots i_d$ for $i \in N^d$ and $i \geq j$ if $i_k \geq j_k$ for all k. By $Q(q,d)$ we mean

$$Q(q,d) = \{i = (i_1, i_2, \ldots i_d) : 1 \leq i, |i| \leq q\}$$

with $1 = (1, 1 \ldots 1)$ and $|Q(q,d)| = \binom{q}{d}$.

We have

$$\begin{aligned} A(q,d) &= \sum_{i \in Q(q,d)} \bigotimes_{k=1}^{d} \Delta_{i_k} = \sum_{i \in Q(q-1,d-1)} (\bigotimes_{k=1}^{d} \Delta_{i_k}) \otimes \sum_{i_d=1}^{g-|i|} \Delta_{i_d} \\ &= \sum_{i \in Q(q-1,d-1)} (\bigotimes_{k=1}^{d-1} \Delta_{i_k}) \otimes U_{q-|i|} \end{aligned} \tag{7}$$

since $\sum_{i=1}^{m} \Delta_i = U_m$ for any $m \geq 1$.

Observe that

$$\bigotimes_{k=1}^{d}(U_{i_k} - U_{i_{k-1}}) = \sum_{\alpha \in \{0,1\}^d} (-1)^{|\alpha|} \bigotimes_{k=1}^{d} U_{i_k - \alpha_k}$$

$\bigotimes_{k=1}^{d} U_{j_k}$ appears in $A(q,d)$ for all indices i for which $i_k = j_k + \alpha_k$ with $\alpha \in \{0,1\}^d$ and $|\alpha| \leq q - |j|$. The sign of $\bigotimes_{k=1}^{d} U_{j_k}$ in this case is $(-1)^{|\alpha|}$.

Let

$$b(i,d) = \sum_{\alpha \in \{0,1\}^d, |\alpha| \leq i} (-1)^{|\alpha|}.$$

This yield

$$A(q,d) = \sum_{j \in Q(q,d)} b(q - |j|, d) \bigotimes_{k=1}^{d} U_{j_k}.$$

We now compute $b(i,d)$. Since $|\alpha| = j$ corresponds to $\binom{d}{j}$ terms, we have

$$b(i,d) = \sum_{j=0}^{\min\{i,d\}} \binom{d}{j} (-1)^j = (-1)^i \binom{d-1}{i}.$$

In particular, $b(i,d) = 0$ for $i \geq d$. This yields the explicit form of $A(q,d)$:

Lema 1.

$$A(q,d) = \sum_{q-d+1 \leq |i| \leq q} (-1)^{q-|i|} \binom{d-1}{q-|i|} (U_{i_1} \otimes \ldots \otimes U_{i_d}) \tag{8}$$

In particular, for

$$U_i(f) = \sum_{j=1}^{m_i} a_{i,j} L_{i,j}(f)$$

with $a_{i,j} \in G_1$ and continuous functionals $L_{i,j}$ we have

$$A(q,d)f = \sum_{q-d+1 \leq |i| \leq q} (-1)^{q-|i|} \binom{d-1}{q-|i|} \sum_{j \leq m_i} L_{i,j}(f) g_{i,j},$$

where $L_{i,j} = \bigotimes_{k=1}^{d} L_{i_k,j_k}$, $g_{i,j} = \bigotimes_{k=1}^{d} a_{i_k,j_k}$ and $m_i = (m_{i_1}, \ldots, m_{i_d})$.

Furthermore we consider the case in which for $d = 1$ we have one of the spaces

$$F_1^r = C^r([-1,1]), \qquad r \in N$$

with the norm

$$||f|| = max(||f||_\infty, \dots, ||f^{(r)}||_\infty).$$

For $d > 1$ consider the tensor product

$$F_d^r = \{f : [-1,1]^d \to \mathbf{R} / D^\alpha f \text{ continuous if } \alpha_i \le r \ \ \forall \ \ i\}$$

with the norm

$$||f|| = max\{||D^\alpha f||_\infty / \alpha \in N_0^d, \alpha_i \le r\}.$$

Let

$$I_d(f) = \int_{[-1,1]^d} f(x)dx, \qquad \text{with } f \in F_d^r. \tag{9}$$

We wish to find good approximation to the functional I_d on the basis of good approximation in the univariate case, using the algorithm of Smolyak.

In the multivariate case $d \ge 1$, define

$$U_{i_1} \otimes \dots \otimes U_{i_d} = \sum_{j_1=1}^{m_{i_1}} \dots \sum_{j_d=1}^{m_{i_d}} f(x_{j_1}^{i_1}, \dots, x_{j_d}^{i_d})(a_{j_1}^{i_1}, \dots, a_{j_d}^{i_d})$$

where we assume that a sequence of quadrature formulas

$$U_i(f) = \sum_{j=1}^{m_i} f(x_j^i) a_j^i$$

is given with $m_i \in N$.

On the basis of Lemma 1 with given quadrature formulas U^i we can write the approximation formula $A(q,d)$ for general d.

$A(q,d)$ is a linear functional, and for $f \in F_d^r$, $A(q,d)(f)$ depends only through function values at a finite number of points.

Let $X^i = \{x_1^i, \dots, x_{m_i}^i\} \subset [-1,1]$ denote the set of points that correspond to U^i. Then $U_{i_1} \otimes \dots \otimes U_{i_d}$ is based on the grid $X^{i_1} \times \dots \times X^{i_d}$, and therefore $A(q,d)(f)$ depends on the values of f at the union

$$H(q,d) = \bigcup_{q-d+1 \le |i| \le q} (X^{i_1} \times \dots \times X^{i_d}) \in [-1,1]^d.$$

If $X_i \subset X_{i+1}$, than $H(q,d) \subset H(q+1,d)$ and $H(q,d) = \bigcup_{|i|=q} (X^{i_1} \times \ldots \times X^{i_d})$. Therefore this kind of sets seems to be the most economical choice.

In the general case we assume that the algorithm

$$U_i(f) = \sum_{j=1}^{m_i} a_{i,j} L_{i,j}(f)$$

use nested information $N_i = [L_{i,1}, L_{i,2}, \ldots, L_{i,m_i}]$. That is,

$$\{L_{i,1}, L_{i,2}, \ldots, L_{i,m_i}\} \subset \{L_{i+1,1}, L_{i+1,2}, \ldots, L_{i+1,m_{i+1}}\}, \ \forall i = 1, 2, \ldots \qquad (10)$$

Since X_1 is now a Hilbert space, $L_{i,j} =< f, f_{i,j} >$ for some element $f_{i,j}$ of X_1. Hence, there exists a sequence $\{f_i\}$ in F_1 such that

$$N_i(f) = \{< f, f_1 >, < f, f_2 >, \ldots, < f, f_m >\} \ \imath = 1, 2, \ldots$$

Assume that the algorithms U_i are optimal, i.e. they minimize the error among all algorithms that use the information N_i. U_i is optimal if

$$L_i = S_1 \mathcal{P}_{\mathrm{i}}, \qquad (11)$$

where $\mathcal{P}$ is the orthogonal projection on the linear subspace $span\{f_j, j = 1, 2, \ldots, m_i\} = (kerN_i)^{\perp}$. Then (11) implies optimality of the algorithm $A(q,d)$ for any d. If we note $N_{q,d}(f) = [L_{i,j}(f) : 1 \leq i, q - d + 1 \leq | i | \leq q, j \leq m_i]$ the information used by the algorithm $A(q,d)$, then for nested information N_i and optimal U_i of (11), $A(q,d) = S_d \mathcal{P}(q,d)$ where $\mathcal{P}(q,d)$ is the orthogonal projection on the linear subspace $(ker(N(q,d))^{\perp}$. Thus, in particular, $A(q,d)$ minimizes the error among all algorithms that use the same information $N_{q,d}$.

4. The Clenshaw-Curtis method

For any cubature formula Q we have the error bound

$$|I_d(f) - Q(f)| \leq ||I_d - Q|| \cdot ||f||$$

In the univariate case $d = 1$

$$\lim_{n \to \infty} n^r \cdot \inf_{Q_n}(||I_1 - Q_n||) = \beta_r \qquad (12)$$

where $\beta_r > 0$ are known constants for any $\forall r \in N$, (Strauβ, 1979), and Q_n are formulas which use n function value.

Novak and Ritter suggest to use the Clenshaw-Curtis method, with a suitable choice of the sequence m_i, where m_i denotes the number of function value used by U_i, and assume that $m_i < m_{i+1}$. In light of (12) they are interested in formulas U_i with

$$\lim_{i\to\infty} \sup(m_i^r \|I_1 - U_i\|) < \infty, \qquad \forall i \in N. \tag{13}$$

and the property is true, for interpolatory formulas U_i, with positive weights.

To obtain nested sets of points, they choose

$$m_i = 2^{i-1} + 1, \quad i > 1 \text{ and } m_1 = 1. \tag{14}$$

Let

$$x_j^i = -\cos\frac{\pi(j-1)}{m_i - 1}, \qquad j = 1, 2, \ldots, m_i$$

and $x_1^1 = 0$, then $U_1(j) = 2j(0)$.

The weights of the Clenshaw-Curtis formula

$$U_i(f) = \sum_{j=1}^{m_i} f(x_j^i) a_j^i$$

are caracterized by the demand that U_i is exact for all polinomials of degree less than m_i, and for $i > 1$ they are given by

$$a_j^i = a_{m_i+1-j}^i = \frac{2}{m_i - 1}\left(1 - \frac{\cos(\pi(j-1)}{m_i(m_i-2} - 2\sum_{k=1}^{m_i-3/2} \frac{1}{4k^2-1}\cdot\cos\frac{2\pi k(j-1)}{m_i - 1}\right)$$

for $j = 2, \ldots, m_i$ and $a_1^i = a_{m_i}^i = \frac{1}{m_i(m_i-2)}$.

For delimitation of the error, they start from the estimate in the univariate case

$$\|I_1 - U_i\| \le \gamma_r \cdot 2^{-r\cdot i}.$$

From (6) we get

$$A(q+1,d+1) = \sum_{|i|\le q}(\Delta^{i_1}\otimes\ldots\otimes\Delta^{i_d}\otimes\sum_{k=1}^{q+1-|i|}\Delta^{i_k})$$
$$= \sum_{|i|\le q}(\Delta^{i_1}\otimes\ldots\otimes\Delta^{i_d}\otimes U_{q+1-|i|}).$$

Then for the error we can obtain the following estimate:

$$I_{d+1} - A(q+1,d+1) = (I_d - A(q,d))\otimes I_1 + \sum_{|i|\le q}\Delta^{i_1}\otimes\ldots\otimes\Delta^{i_d}\otimes(I_1 - U_{q+1-|i|}).$$

Furthermore

$$||\Delta^{i_k} \le ||I_1 - U_{i_k}|| + ||I_1 - U_{i_k-1}|| \le \gamma_r\cdot 2^{-ri_k}(1+2^r).$$

We get

$$\sum_{|i|\le q}||\Delta^{i_1}||\cdot\ldots\cdot||\Delta^{i_d}||\cdot||I_1 - U^{q+1-|i|}|| \le \binom{q}{d}\cdot\gamma_r^{d+1}\cdot(1+2^r)^d\cdot 2^{-r(q+1)}.$$

Inductively the following theorem can be obtained.

Theorem 1. *Let $\theta_r = \max\{2^{r+1}, \gamma_r\cdot(1+2^r)\}$. The error of the cubature formula $A(q,d)$ satisfies the following estimates:*

$$||I_d - A(q,d)|| \le \gamma_r\theta_r^{d-1}\binom{q}{d-1}\cdot 2^{-r\cdot q}.$$

Corollary 1. *Let $n = n(q,d)$ denote the number of knots used by $A(q,d)$. Then*

$$||I_d - A(q,d)|| = \mathcal{O}(n^{-r}\cdot(\log n)^{(d-1)(r-1)}).$$

This corrolary gives the error of $A(q,d)$ related to the number of knots from $H(q,d)$ and also gives the best error bound for Smolyak's algorithm which holds for arbitrary tensor product problems. On the other hand this method yields error of order $n^r(logn)^{(d-1)(r-1)}$ for all classes F_d^r, hence this methods are almost optimal up to logarithmic factors on a whole scale of spaces of nonperiodic functions.

Property (15) is the essential requirement for the U_i in the univariate case. Relation which also holds for the Gauss formulas. These formulas yield methods

$A(q,d)$ with a higher degree of exactness. Still Novak and Ritter prefer the Clenshaw-Curtis formulas because in this case the number of knots from $H(q,d)$ is reduced. Weights of different signs at common points are partially cancelled.

To determine the polynomial exactness they start from the fact that the Clenshaw-Curtis formula U_i is exact on $V^i = P_{m_i}$, where m_i is odd.

Theorem 2. *The cubature formula $A(q,d)$ is exact on*

$$\sum_{|i|=q} (V^{i_1} \otimes \ldots \otimes V^{i_d}).$$

The theorem can be proved by induction over d.

Remark. Theorem 2 holds for general tensor product problems if the space

$$V^i = \{f \in F_1^r / I_1(f) = U^i(f)\}$$

of exactness for the univariate problem is nested, $V^i \subset V^{i+1}$.

References

[1] Brass, H.,*Error bounds based on approximation theory*, in Espelid, T.O.,Genz,A.,eds., Numerical Integration, pp.147-163. Kluwer Academic Publishers,Dordrecht, 1992.
[2] Delvos, F.-J., *d-variate Boolean interpolation*, J.Approx.Th. **34** (1982), 99-114
[3] Delvos, F.-F., *Boolean methods for double integration*, Math. of Comp., **55** (1990),683-692
[4] Engels, H., *Numerical Quadrature and Cubature*, Academic Press, London, 1980
[5] Novak, E., Ritter, K., *High-dimensional integration of smooth functions over cubes*, Numer. Math., **75** (1996), 79-97
[6] Smolyak, S.A., *Quadrature and interpolation formulas for tensor products of certain classes of function*, Soviet Math. Dokl., 4 (1963), 240-243
[7] Strauβ, H., *Optimal quadrature formulas*, in Meinardus,G., Approximation in Theory und Praxis, pp. 239-250, Bibliographisches Institut, Mannheim, 1979
[8] Temlyakov, V.N., *On approximate recovery of functions with bounded mixed derivative*, J. Complexity, **9** (1987), 41-59

"BABEŞ-BOLYAI" UNIVERSITY, FACULTY OF MATHEMATICS AND COMPUTER SCIENCE, 3400 CLUJ-NAPOCA, ROMANIA.