Interner Bericht

Complexity of Multivariate Integral Equations: Full Solution in Sobolev Spaces

Karin Frank

271/95

Fachbereich Informatik

Universität Kaiserslautern · Postfach 3049 · D-67653 Kaiserslautern

Complexity of Multivariate Integral Equations: Full Solution in Sobolev Spaces

Karin Frank

271/95

Universität Kaiserslautern
AG Numerische Algorithmen
Postfach 30 49
67653 Kaiserslautern

Mai 1995

Herausgeber: AG Numerische Algorithmen
Leiter: Professor Dr. S. Heinrich

Complexity of Multivariate Integral Equations: Full Solution in Sobolev Spaces

Karin Frank*

Fachbereich Informatik, Universität Kaiserslautern
Postfach 3049, D–67653 Kaiserslautern

17th May 1995

Abstract

The problem of full solution of Fredholm integral equations of the second kind with data from Sobolev spaces with dominating mixed derivative is studied. Existing estimates for the univariate case are extended to arbitrary dimension.

1 Introduction

Information based complexity theory (IBC) studies the essential difficulty of a numerical problem. That means it searches the answer of the question, what is the minimal cardinality of information about the input data and the minimal number of arithmetical operations required to solve the numerical problem with an error less than ε. In this paper, we consider the problem of numerical solution of Fredholm integral equations of the second kind with kernels and free terms belonging to Sobolev spaces with dominating mixed derivatives.

The first result on Fredholm integral equations in the setting of IBC was obtained by Emelyanov and Ilin [EI67]. They restricted themselves to the study of algorithms using as information functionals only function values of the kernel and the free term at some points, assuming both functions to be r–times continuously differentiable.

Integral equations in Sobolev classes with dominating mixed derivative were treated by Pereverzev in [Per91], permitting as information functionals arbitrary linear

*e-mail: frank@informatik.uni-kl.de

functionals. He derived lower and upper estimates of the information complexity in the onedimensional case. In the present paper, we will extend these results to the multivariate case with arbitrary dimension d.

The paper leads out as follows: In section 2, the problem is formulated. All necessary definitions are outlined, for more detail the reader is referred to [TWW88]. Furthermore, in Theorem 1 the main result is stated. It represents the exact order of the so-called n-th minimal radius of information, a quantity closely related to the ε-complexity of a numerical problem. In section 3, the lower bound of Theorem 1 is shown, using a Gelfand number technique introduced by Heinrich [Hei93] and generalized by Frank, Heinrich and Pereverzev in [FHP95]. The proof of the upper bound is the subject of section 4, where an optimal algorithm is described and analyzed. This algorithm performs a two-level iteration, where the finer level represents a hyperbolic cross type approximation of the kernel function. Hyperbolic crosses were first considered by Babenko [Bab60] for function approximation in Sobolev spaces with dominating mixed derivative.

2 Problem setting

Let $G = [0,1]^d$ with $d \in \mathbb{N}$, and $L_2(G)$ be the space of square summable functions on G. We consider the orthonormal trigonometric basis in $L_2([0,1])$

$$\begin{aligned} e_0(\tau) &\equiv 1 \\ e_n(\tau) &= \sqrt{2}\cdot\cos(2\pi n\tau) \\ e_{-n}(\tau) &= \sqrt{2}\cdot\sin(2\pi n\tau) \end{aligned}$$

for arbitrary $n \in \mathbb{N}$. Then for a given multiindex $i = (i_1, \ldots, i_d) \in \mathbb{Z}^d$ the basis function $e_i \in L_2(G)$ is defined by

$$e_i(t) = e_{i_1}(t_1)\cdot\ldots\cdot e_{i_d}(t_d) \qquad (t = (t_1, \ldots, t_d) \in G).$$

The Fourier coefficients of $f \in L_2(G)$ are given by

$$\hat{f}(i) = (f, e_i) \qquad (i \in \mathbb{Z}^d).$$

Similarly, an orthonormal basis $\{e_{ij}\}_{i,j\in Z^d}$ in $L_2(G^2)$ is defined by

$$e_{ij}(s,t) = e_i(s)\cdot e_j(t) \qquad (s,t \in G).$$

Then, the Fourier coefficients of $k \in L_2(G^2)$ are of the following form

$$\hat{k}(i,j) = (k, e_{ij}) \qquad (i,j \in \mathbb{Z}^d).$$

Now we shall define the class of data to be discussed. Therefore, given a multiindex $i = (i_1, i_2, \dots, i_d) \in \mathbb{Z}^d$ we set $\rho(i) = \max(1, |i_1|) \cdot \max(1, |i_2|) \cdot \ldots \cdot \max(1, |i_d|)$, where $|i_k|$ denotes the absolute value of $i_k \in \mathbb{Z}$. Let $r \geq 0$. Then the function spaces $\mathcal{H}^r(G)$ and $\mathcal{H}^r(G^2)$ are defined as

$$\begin{aligned}
\mathcal{H}^r(G) &= \{f \in L_2(G) : \|f\|_r^2 = \sum_{i \in Z^d} \rho(i)^{2r} \hat{f}(i)^2 < \infty\}, \\
\mathcal{H}^r(G^2) &= \{k \in L_2(G^2) : \|k\|_{r,r}^2 = \sum_{i,j \in Z^d} \rho(i)^{2r} \rho(j)^{2r} \hat{k}(i,j)^2 < \infty\}.
\end{aligned}$$

For simplicity, we will often use the following notation: $\mathcal{H}^r = \mathcal{H}^r(G)$, $\mathcal{H}^{r,r} = \mathcal{H}^r(G^2)$, $L_2 = L_2(G)$. Note that for $r \in \mathbb{N}$ the space $\mathcal{H}^r(G)$ constitutes the Sobolev space of periodic functions f on $[0,1]^d$, for which both f and the generalized mixed derivative $\frac{\partial^{dr} f}{\partial^r t_1 \ldots \partial^r t_d}$ belong to L_2. These spaces are called Sobolev spaces with dominating mixed derivative.

By $\mathcal{H}^{-r} = (\mathcal{H}^r)^*$ we denote the dual space of $\mathcal{H}^r$. L_2 imbeds into $\mathcal{H}^{-r}$ in a canonical way, and the $\mathcal{H}^{-r}$-norm of a function $f \in L_2$ is given by

$$\|f\|_{-r}^2 = \sum_{i \in Z^d} \rho(i)^{-2r} \hat{f}(i)^2 .$$

Note that L_2 is a dense subspace of $\mathcal{H}^{-r}$.

Finally, we define subsets $F_0 \subset \mathcal{H}^r(G)$, $K_0 \subset \mathcal{H}^r(G^2)$ of the form

$$\begin{aligned}
F_0 &= \{f \in \mathcal{H}^r(G) : \|f\|_r \leq \gamma\}, \\
K_0 &= \{k \in \mathcal{H}^r(G^2) : \|k\|_{r,r} \leq \alpha, \|(I - T_k)^{-1} : L_2 \to L_2\| \leq \beta\},
\end{aligned}$$

where $\alpha, \gamma > 0$ and $\beta > 1$.

Now we are ready to state the problem to be studied. We consider integral equations of the form

$$u - T_k u = f, \qquad (1)$$

where $f \in F_0$, $k \in K_0$, and T_k denotes the integral operator

$$\begin{aligned}
T_k &: L_2(G) \to L_2(G) \\
T_k u &= \int_G k(s,t)\, u(t)\, dt .
\end{aligned}$$

The problem is to be formulated within the framework of information–based complexity theory. Here only the most important definitions are outlined, referring to [TWW88] for further notations.

Since we are interested in the solution of (1) on the whole domain G, we have to consider the so-called full solution operator

$$\begin{aligned} S &: K_0 \times F_0 \rightarrow L_2(G) \\ S(k,f) &= (I - T_k)^{-1} f. \end{aligned}$$

This operator describes the exact solution of the numerical problem we have to solve approximately by means of some partial information about the concrete given equation. We permit linear information on the kernel function and the right-hand side, i.e. the information operator is defined by $N : K_0 \times F_0 \rightarrow \mathbb{R}^n$, $N = (N_1, N_2)$ with

$$\begin{aligned} N_1 k &= ((k, g_1), \ldots, (k, g_{n_1})) \ , \ g_k \in \mathcal{H}^{-r}(G^2) \quad (k = 1, \ldots, n_1) \\ N_2 f &= ((f, h_1), \ldots, (f, h_{n_2})) \ , \ h_l \in \mathcal{H}^{-r}(G) \quad (l = 1, \ldots, n_2) \end{aligned}$$

where $n_1 + n_2 = n$.

An approximation to the exact solution $S(k,f)$ is to be computed. An arbitrary mapping $\varphi : \mathbb{R}^n \rightarrow L_2(G)$, which combines the information $N(k,f)$ and computes an approximation $\varphi(N(k,f))$ to S, is called an algorithm. Then the error of an approximation $\varphi(N(k,f))$ is defined by

$$e(S, N, \varphi) = \sup_{f \in F_0, k \in K_0} \|S(k,f) - \varphi(N(k,f))\|_{L_2} .$$

Let us agree about the model of computation. We assume, that standard arithmetical operations, including comparisons, can be performed with unit cost, while linear functionals on the input data can be computed with constant cost $c(d)$. Imagine a subroutine which supplies the computation of one linear functional on the data.

2.1 The main result

Our main theorem provides estimates for the n-th minimal radius of information of the given problem. This quantity describes the minimal error, which can be obtained by any algorithm φ using at most n information functionals:

$$e_n(S) = \inf_{N(k,f) \in R^n} \inf_{\varphi : R^n \rightarrow L_2} e(S, N, \varphi)$$

This is the crucial quantity to be analyzed in information-based complexity. Since any algorithm of cost n can use at most n information functionals due to the model of computation, $e_n(S)$ serves as a general lower bound for the error of any algorithm of cost n.

Theorem 1 *Let $r > 0$. Then there exist constants $c_1, c_2 > 0$ such that for all $n \in \mathbb{N}$*

$$c_1 \cdot n^{-r} \log^{r(d-1)} n \ \le \ e_n(S) \ \le \ c_2 \cdot n^{-r} \log^{r(d-1)} n .$$

3 Proof of the lower bound

Let us define some mapping Ψ by

$$\begin{aligned} \Psi &: \mathcal{H}^r(G^2) \to L(\mathcal{H}^r(G), L_2(G)) \\ \Psi k &= T_k : \mathcal{H}^r(G) \to L_2(G) \end{aligned}$$

and introduce the so-called Gelfand numbers of an operator. Given two Banach spaces E and F, let B_E denote the unit ball of E and $L(E,F)$ the space of all bounded linear operators from E to F. Then for an operator $T \in L(E,F)$ and $n \in \mathbb{N}$ the n-th Gelfand number of T is defined by

$$c_n(T) = \inf_{\lambda_1,\ldots,\lambda_{n-1}\in E^*} \sup_{\substack{x\in B_E \\ \lambda_1(x)=\ldots=\lambda_{n-1}(x)=0}} \|Tx\| .$$

For details on these numbers we refer to [Pie78]. Using Theorem 1 by [FHP95] we know that

$$e_n(S) \succ c_n(\Psi : \mathcal{H}^r(G^2) \to L(\mathcal{H}^r(G), L_2(G))) , \tag{2}$$

To prove the lower bound of Theorem 1, we are going to estimate the Gelfand numbers of the operator Ψ. But first we provide a lemma which we will use afterwards.

Lemma 1 *Let $\Gamma_n^{(d)}$ be the set of multiindices*

$$\Gamma_n^{(d)} = \{(i,j) \in \mathbb{Z}^{2d} : \rho(i)\cdot\rho(j)^2 \le n\} .$$

Then for any $d \in \mathbb{N}$ the cardinality of the set $\Gamma_n^{(d)}$ is

$$|\Gamma_n^{(d)}| \asymp n \cdot \log^{(d-1)} n$$

Proof:

We will show the lemma by induction over the dimension d. For $d = 1$ we get by elementary integration:

$$\begin{aligned} |\Gamma_n^{(1)}| &\asymp \sum_{i_1=1}^{n} \frac{\sqrt{n}}{\sqrt{i_1}} \\ &\asymp \int_1^n \frac{\sqrt{n}}{\sqrt{i_1}}\, di_1 \\ &= \sqrt{n}\cdot 2(\sqrt{n}-1) \asymp n . \end{aligned}$$

Now we prove the statement of the lemma for any $d \in \mathbb{N}$ under the assumption, that it is already shown for the dimension $d-1$, that means

$$|\Gamma_n^{(d-1)}| \asymp n \cdot \log^{(d-2)} n \,.$$

Then we get

$$\begin{aligned}
|\Gamma_n^{(d)}| &\asymp \sum_{i=1}^{n} \sum_{j=1}^{\frac{\sqrt{n}}{\sqrt{i}}} \left| \Gamma^{(d-1)}_{[\frac{n}{ij^2}]} \right| \\
&\asymp \int_1^n \int_1^{\frac{\sqrt{n}}{\sqrt{i}}} \frac{n \cdot \log^{d-2}\left(\frac{n}{ij^2}\right)}{ij^2}\, dj\, di \\
&= \int_1^n \int_{\frac{\sqrt{n}}{\sqrt{i}}}^1 \sqrt{\frac{n}{i}} \cdot (-\log^{d-2} x^2)\, dx\, di \\
&\asymp \int_1^n \frac{n}{i} \cdot \log^{d-2} \sqrt{\frac{n}{i}}\, di \\
&\asymp \int_{\log n}^0 -n \cdot y^{d-2}\, dy \\
&\asymp n \cdot \log^{d-1} n \,,
\end{aligned}$$

where we used the substitutions $x = \frac{\sqrt{n}}{\sqrt{i \cdot j}}$, $y = \log \sqrt{\frac{n}{i}}$, and the fact that for $s \in \mathbb{N}$ and $m > 1$

$$\begin{aligned}
\int_1^m \log^s z\, dz &= z \cdot \sum_{k=1}^{s} \frac{(-1)^{s-k} s!}{k!} \log^k z \Bigg|_1^m \\
&\asymp m \cdot log^s m \,.
\end{aligned}$$

□

Let $\{b_{ij}\}_{i,j \in Z^d}$ be the unit vector basis in $l_2(\mathbb{Z}^{2d})$. Then we define two operators

$$\begin{aligned}
W &: l_2(\mathbb{Z}^{2d}) \to \mathcal{H}^{r,r} \\
W b_{ij} &= \rho(i)^{-r} \cdot \rho(j)^{-r} \cdot e_i(t) \cdot e_j(s) \,, \\
U &: L(\mathcal{H}^r, L_2) \to l_\infty(\mathbb{Z}^{2d}) \\
UA &= \rho(m)^{-r} \cdot (Ae_m, e_l) \,,\ m, l \in \mathbb{Z}^d \,.
\end{aligned}$$

As easily can be seen, the operator W is an isometry, so $\|W\| = 1$, whereas the operator U is an injection with $\|U\| \le 1$. Composing these two operators with Ψ we define

$$\begin{aligned}
D &: l_2(\mathbb{Z}^{2d}) \to l_\infty(\mathbb{Z}^{2d}) \\
D &= U \Psi W \,.
\end{aligned}$$

The operator D is a diagonal operator:

$$\begin{aligned} D\, b_{ij} &= \xi_{ij} \cdot b_{ij} \\ \xi_{ij} &= \rho(i)^{-r} \cdot \rho(j)^{-2r} . \end{aligned}$$

Let $\{\lambda_k\}_{k\in N}$ be a rearrangement of the numbers $\{\xi_{ij}\}_{i,j\in Z^d}$ in nonincreasing order. Then we have the following

$$\begin{aligned} \lambda_n &= \inf\left\{\varepsilon : \mathrm{card}\{(i,j) \in \mathbb{Z}^{2d} : \xi_{ij} > \varepsilon\} < n\right\} \\ &= \max_{\substack{\Gamma\subseteq Z^{2d} \\ |\Gamma|=n}} \min\{\xi_{ij} : (i,j) \in \Gamma\} . \end{aligned} \tag{3}$$

Let $\Gamma_n = \{(i,j) \in \mathbb{Z}^{2d} : \xi_{ij} \geq n^{-r}\}$. Lemma 1 gives us

$$\mathrm{card}\,(\Gamma_n) \asymp n \cdot \log^{(d-1)} n .$$

From (3) it follows that

$$\begin{aligned} \lambda_{[\mathrm{card}(\Gamma_n)]} &\geq n^{-r} > \lambda_{[\mathrm{card}(\Gamma_n)+1]} \\ \lambda_n &\asymp n^{-r} \cdot \log^{r(d-1)} n . \end{aligned}$$

Then it follows from Theorem 11.11.7 in [Pie78] that

$$c_n(D : l_2(\mathbb{Z}^{2d}) \to l_\infty(\mathbb{Z}^{2d})) \asymp \lambda_n \asymp n^{-r} \cdot \log^{r(d-1)} n .$$

By basic properties of Gelfand numbers we conclude

$$\begin{aligned} c_n(\Psi : \mathcal{H}^{r,r} \to L(\mathcal{H}^r, L_2)) &\geq \|U\| \cdot c_n(\Psi : \mathcal{H}^{r,r} \to L(\mathcal{H}^r, L_2)) \cdot \|W\| \\ &\geq c_n(D : l_2(\mathbb{Z}^{2d}) \to l_\infty(\mathbb{Z}^{2d})) \\ &\succ n^{-r} \cdot \log^{r(d-1)} n . \end{aligned}$$

Together with (2) this proves the lower bound of Theorem 1. □

4 The optimal algorithm

In [Per91] an optimal algorithm for dimension $d = 1$ was proposed. This direct algorithm uses a hyperbolic cross type approximation of the kernel function. The number of operations was reduced by representing the right–hand side and the approximate solution not in the original basis, but in another generating system. We did not try to extend this approach to the multivariate case. Our algorithm uses a similar approach to that in [FH94], [Fra94] and [FHP95], that means we perform a two–level iteration, solving a system of linear algebraic equations only on the coarser level and just summing up the remaining coefficients on the finer level, which corresponds to the hyperbolic cross approximation of the kernel function.

Let $k \in K_0$, $f \in F_0$ be given, $n \in \mathbb{N}$ arbitrary, but fixed. Then let us define the index sets

$$\begin{aligned}
\Gamma_n &= \left\{(i,j) \in \mathbb{Z}^{2d} \, : \, \rho(i)\cdot\rho(j)^2 \leq n\right\} \\
A_n &= \left\{i \in \mathbb{Z}^{d} \, : \, \rho(i) \leq n^{\frac{1}{4}}\right\} \\
B_n &= \left\{i \in \mathbb{Z}^{d} \, : \, \rho(i) \leq n\right\} \\
C_n &= \left\{(i,j) \in \mathbb{Z}^{2d} \, : \, \max\left\{\rho(i),\rho(j)\right\} \leq n^{\frac{1}{4}}\right\} .
\end{aligned}$$

Note that the set Γ_n is the same as in section 3. The cardinalities of these sets are

$$\begin{aligned}
|\Gamma_n| &\asymp n \cdot \log^{d-1} n \\
|A_n| &\asymp n^{\frac{1}{4}} \cdot \log^{d-1} n \\
|B_n| &\asymp n \cdot \log^{d-1} n \\
|C_n| &\asymp n^{\frac{1}{2}} \cdot \log^{2(d-1)} n .
\end{aligned}$$

This follows from Lemma 1 for Γ_n and can be verified for the other sets by similar, but easier calculations. The projections g, h and f_0 of k and f, respectively, are defined by its Fourier coefficients

$$\begin{aligned}
f_0 \in \mathcal{H}^r(G) \; &: \; \hat{f}_0(i) = \begin{cases} \hat{f}(i) : i \in B_n \\ \;0 \;\; : \text{otherwise}, \end{cases} \\
g \in \mathcal{H}^r(G^2) \; &: \; \hat{g}(i,j) = \begin{cases} \hat{k}(i,j) : (i,j) \in C_n \\ \;\;0 \;\;\; : \text{otherwise}, \end{cases} \\
h \in \mathcal{H}^r(G^2) \; &: \; \hat{h}(i,j) = \begin{cases} \hat{k}(i,j) : (i,j) \in \Gamma_n \\ \;\;0 \;\;\; : \text{otherwise}. \end{cases}
\end{aligned}$$

The iteration process is described by the following equation

$$(I - T_g)\, v_l = f_0 + (T_h - T_g)\, v_{l-1} \tag{4}$$

for $l = 1, \ldots, 4$, $v_0 \equiv 0$. The number of iterations is determined by the norm of the operator $(T_h - T_g) : L_2 \to L_2$, as follows from Lemma 2(iii) and Corollary 1 below. In terms of Fourier coefficients, we have to solve on each iteration step

$$i \in A_n \; : \; \hat{v}_l(i) - \sum_{j \in A_n} \hat{k}(i,j)\cdot\hat{v}_l(j) = \hat{f}(i) + \sum_{j:(i,j)\in\Gamma_n\setminus C_n} \hat{k}(i,j)\cdot\hat{v}_{l-1}(j) , \tag{5}$$

$$i \in B_n \setminus A_n \; : \; \hat{v}_l(i) = \hat{f}(i) + \sum_{j:(i,j)\in\Gamma_n} \hat{k}(i,j)\cdot\hat{v}_{l-1}(j) . \tag{6}$$

The algorithm $\varphi : \mathbb{R}^m \to L_2(G)$ of approximate solution of equation (1) is defined by

$$\varphi(N(k,f)) = v_4(t),$$

where $N : (K_0, F_0) \to \mathbb{R}^m$ is the information operator

$$\begin{aligned} N(k,f) &= (N_1 k, N_2 f), \\ N_1 k &= \left(\hat{k}(i,j)\right)_{(i,j)\in\Gamma_n}, \\ N_2 f &= \left(\hat{f}(i)\right)_{i\in B_n}. \end{aligned}$$

Hence, the cardinality of information is

$$\begin{aligned} m &= \mathrm{card}(\Gamma_n) + \mathrm{card}(B_n) \\ &\asymp n \cdot \log^{d-1} n . \end{aligned}$$

For the correctness of the algorithm we have to say that equation (4) is uniquely solvable in L_2, as follows from Lemma 4(ii) below. Equation (4) is equivalent to equations (5) and (6), since $\hat{v}_l(i) = 0$ for all $i \notin B_n$, as a consequence of $\hat{f}(i) = \hat{g}(i,j) = \hat{h}(i,j) = 0$ for all $i, j \in \mathbb{Z}^d$ with $\rho(i) > n$.

Now we have to estimate the number of arithmetical operations required. On each iteration step, the matrix of the system (5) of linear algebraic equations can be computed in $O(|\Gamma_n|)$ operations, the system itself can be solved in $O(|A_n|^3)$ operations, e.g. by Gaussian elimination. The summation in (6) can be performed in $O(|\Gamma_n|)$ operations. In total, we have to do on each iteration step $O(n \cdot \log^{d-1} n)$ arithmetical operations.

With $\mathrm{cost}(\varphi)$ and $\mathrm{card}(N)$ both of the order $O(n \cdot \log^{d-1} n)$, our algorithm will be optimal, if its error satisfies

$$e(S, N, \varphi) \prec n^{-r} . \qquad (7)$$

To prove this relation, we shall rewrite the algorithm in a more convenient form and provide three lemmas. The proofs of the lemmas will be dropped or given in a very short form, because they are similar to those in [Fra94]. Let

$$\begin{aligned} Y &= (I - T_g)^{-1} (T_h - T_g) \\ Z &= (I - T_g)^{-1} (T_k - T_h) \\ w &= (I - T_k)^{-1} f_0 . \end{aligned}$$