ON THE EXPANSION OF FUNCTIONS IN APPROXIMATELY ORTHONORMAL SYSTEMS

MATHEMATICS

1969

SovietRxiv

Abstract

Full Text

UDC 517.9

MATHEMATICS

M. A. ALEKSIDZE, N. M. ARVELADZE, N. L. LEKISHVILI, K. V. PERTAIA

ON THE EXPANSION OF FUNCTIONS IN APPROXIMATELY ORTHONORMAL SYSTEMS

(Presented by Academician I. N. Vekua on 30 IX 1968)

Let $\{\omega_k\}$ be a linearly independent and complete system of functions in L_2. In order to find the coefficients a_k^n of the expansion of a function $\gamma \subset L_2$ in the series

$$\sum_{k=1}^{n} a_k^{(n)} \omega_k,$$

usually from the system $\{\omega_k\}$ one obtains an orthonormalized system

$$\{\psi_k\} \left(\psi_k = \sum_{i=1}^{k} A_{k,i} \omega_i \right),$$

forms for the function γ the Fourier series

$$\sum_{k=1}^{n} b_k \psi_k,$$

where $b_k = (\gamma, \psi_k)$, and then finds a_k^n from the equalities

$$a_k^{(n)} = \sum_{i=k}^{n} A_{i,k} b_i.$$

However, the numerical realization ([1],[2]) of the orthonormalization process for nonminimal systems ([3]) encounters difficulties, and if it is nevertheless possible to realize the corresponding computational scheme, then instead of an orthonormalized system $\{\psi_k\}$ we obtain an approximately orthonormalized system $\{\varphi_k\}$, i.e.

$$(\varphi_k, \varphi_j) = \delta_{k,j} + \varepsilon_{k,j},$$

where $\delta_{k,j}$ is the Kronecker symbol, and $\varepsilon_{k,j}$ are quantities small in absolute value.

It is natural to require of the expansion method that, for an approximately orthonormalized system, the condition

$$\left\|\gamma-\sum_{k=1}^{n}a_k^{(n)}\varphi_k\right\|\leq\left\|\gamma-\sum_{k=1}^{m}a_k^{(m)}\varphi_k\right\|\quad\text{for } n>m, \tag{1}$$

be satisfied, i.e. that increasing the terms of the expansion should not increase the error of the expansion. It is obvious that in the general case for Fourier series condition (1) is not satisfied. The requirement of convergence

$$\left\|\gamma-\sum_{k=1}^{n}a_k^{(n)}\varphi_k\right\|_{L_2}\to 0\quad\text{as } n\to\infty$$

is, generally speaking, without foundation, since even if the systems $\{\psi_k\}$ and $\{\varphi_k\}$ are quadratically close, i.e.

$$\sum_{k=1}^{\infty}\|\psi_k-\varphi_k\|_{L_2}<\infty,$$

then for completeness of the system $\{\varphi_k\}$ it is necessary and sufficient ([3]) that the system $\{\varphi_k\}$ be minimal, which, naturally, cannot be required under a random distribution of the rounding error and of the error of approximate integration (in computing scalar products).

In ([1],[2]) the following method is proposed for finding the coefficients of the expansion of a function γ in the series

$$\sum_{k=1}^{\infty}c_k\varphi_k$$

with respect to an arbitrary system $\{\varphi_k\}$:

$$c_1=(\gamma,\varphi_1)/(\varphi_1,\varphi_1),\qquad c_k=\left(\gamma-\sum_{i=1}^{k-1}c_i\varphi_i,\varphi_k\right)/(\varphi_k,\varphi_k), \tag{2}$$

and sufficient conditions for the convergence of such series are given. Since c_k is—

is the coefficient of the best, in the sense of the L_2 metric, approximation of the function

$$\left(\gamma-\sum_{i=1}^{k-1}c_i\varphi_i\right)$$

by means of the functions φ_k, then it is clear that for this method inequality (1) will always be satisfied.

In the present note it is shown that method (2) for an approximately orthonormalized system gives, in a certain sense, expansions more accurate than the Fourier-series method. The results of numerical calculations given at the end of the note turned out to be unexpected for the authors.

The coefficients λ_k $(k = 1, \dots, n)$ of the best approximation of the function γ by means of the first n terms of the system $\{\varphi_k\}$ are found [4] from the system

$$\sum_{k=1}^{n} \lambda_k(\varphi_j, \varphi_k) = (\gamma, \varphi_j) \qquad (j = 1, 2, \dots, n). \tag{3}$$

The latter system can be solved by iterative methods. The computation of the coefficients of the Fourier series with respect to the system $\{\varphi_i\}$ corresponds to one simple iteration for system (3), when the zero vector $\lambda^{(0)}(0, \dots, 0)$ is taken as the initial approximation for the vector $(\lambda_1, \dots, \lambda_n)$, i.e.

$$a_1^{(\Phi)} = \lambda^{(1)} = B\lambda^{(0)} + R,$$

where $a_1^{(\Phi)}$ is the vector whose components are the Fourier coefficients; B is the matrix corresponding to the writing of system (3) in the form

$$\lambda = B\lambda + R, \tag{4}$$

$$\lambda^{(1)} = (\lambda_1^{(1)}, \dots, \lambda_n^{(1)}), \qquad \lambda^{(0)} = (0, \dots, 0),$$

$$R = \{(\varphi, \varphi_1)/(\varphi_1, \varphi_1), \dots, (\varphi, \varphi_n)/(\varphi_n, \varphi_n)\}.$$

Let us write (4) in the form

$$\lambda = (B_1 + B_2)\lambda + R,$$

where

$$B_1 = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ -(\varphi_1, \varphi_2) & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -(\varphi_1, \varphi_n) & -(\varphi_2, \varphi_n) & \dots & -(\varphi_{n-1}, \varphi_n) & 0 \end{pmatrix}$$

$$B_2 = \begin{pmatrix} 0 & -(\varphi_2, \varphi_1) & -(\varphi_3, \varphi_1) & \dots & -(\varphi_n, \varphi_1) \\ 0 & 0 & -(\varphi_3, \varphi_2) & \dots & -(\varphi_n, \varphi_2) \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

The computation of the coefficients (2) with respect to the system $\{\varphi_i\}$ corresponds to one iteration by the Seidel method for system (4), or else to one simple iteration for the system

$$\lambda = (I - B_1)^{-1} B_2 \lambda + (I - B_1)^{-1} R, \tag{5}$$

where I is the identity matrix, when the zero vector $\lambda^{(0)} = (0, \dots, 0)$ is taken as the initial approximation, i.e.

$$a^{(A)} = \lambda_1^{(1)} = (I - B_1)^{-1} B_2 \lambda^{(0)} + (I - B_1)^{-1} R,$$

where $a^{(A)}$ is the vector whose components are the coefficients of the proposed series. We note that all eigenvalues of the matrix $(I - B_1)^{-1} B_2$ are less than one. Indeed, since the quadratic form corresponding to the Gram matrix is positive definite, then

the Seidel method for system (4), or, what is the same thing, the method of simple iteration for system (5), converges* for any initial approximation and right-hand side.

With such an approach, the theorem proved in (2) is equivalent to the following assertion. For an approximate solution of system (3), for sufficiently large n, a single iteration is sufficient (it is assumed that the system $\{\varphi_i\}$ and the function $\varphi(x)$ satisfy the corresponding conditions), by the Seidel method.

Since the computation of the elements of the Gram matrix (scalar products) and the orthonormalization process itself are carried out with a finite number of digits, it is clear that the matrix corresponding to system (3) will not be the identity matrix. It will be "perturbed" in a certain way. The coefficients of the best approximation are then found from the system

$$(1 - \varepsilon_{jj})\lambda_j + \sum_{k=1,\ k \neq j}^{n} \varepsilon_{k,j} \lambda_k = (\varphi, \varphi_j) \qquad (j = 1, 2, \dots, n), \tag{6}$$

where $\varepsilon_{i,j}$ are small perturbations caused by rounding errors and by the inaccuracy of computing scalar products (in the case of the space L_2, by the integration error). Naturally, one must assume here that

$$\max_k \sum_{j=1}^{n} |\varepsilon_{k,j}| < 1.$$

It is of interest which series is more advisable to use in such cases—the Fourier series or the proposed series? More precisely, what relation exists between the numbers $\|\lambda - a^{(\Phi)}\|$ and $\|\lambda - a^{(A)}\|$, where the vector λ is the exact solution of system (6)?

Since in the case considered the Seidel iteration process converges faster (5) than the simple iteration process, and the matrix norm

$$\max_i \sum_{j=1}^{n} |\varepsilon_{k,j}|$$

is subordinate to the first vector norm, it is clear that the first norm (5) of the vector $(\lambda - a^{(\Phi)})$ will be no smaller than the first norm of the vector $(\lambda - a^{(A)})$, i.e.

$$\max_i |\lambda_i - a_i^{(\Phi)}| = \|\lambda - a^{(\Phi)}\|_I \geq \|\lambda - a^{(A)}\|_I = \max_i |\lambda_i - a_i^{(A)}|.$$

We note that both the Fourier series method and the method of the proposed series require, in order to compute the coefficients of each new term of the expansion, carrying out one approximate integration

$$a_k^{(\Phi)} = \int_G \varphi\varphi_k \, dx, \qquad a_k^{(A)} = \int_G \varphi^{(k-1)}\varphi_k \, dx.$$

In the computation of $a_k^{(A)}$, $\varphi^{(k-1)}$ is involved; therefore it is natural that its values at individual points should be used to control the accuracy of the expansion, since

$$\varphi^{(k-1)}(x) = \varphi(x) - \sum_{i=1}^{k-1} a_i^{(A)}\varphi_i(x).$$

Numerical example. The interior Dirichlet problem was solved:

$$\Delta u = 0, \qquad u|_\Gamma = \operatorname{arc\,tg}(y-2)/(x-2),$$

---

* Since the diagonal elements of the Gram matrix of a linearly independent system are positive, the positive definiteness of the corresponding quadratic form is not only sufficient, but also necessary for the convergence of the Seidel iteration process (E. Reich, see (5)).

where Γ is an ellipse with semiaxes $a = 1,\ b = 0.75$, centered at the origin. The solution was found by means of the standard program ([6]) by the method of functional equations. The auxiliary contour ([8], [9]) was a confocal ellipse with semiaxes $a = 1.7333,\ b = 1.3$.

In Table 1, at the nodes of a grid with step 0.2, the errors of the method of generalized series using Fourier series are given (the upper left number, multiplied by

Table 1

[[unclear: the visible "Table 1" is a plotted elliptical grid with coordinate axes and handwritten numerical entries in each grid cell; most cell entries are not reliably legible from the page image for faithful Markdown transcription.]]

10^{-4}) and using the series proposed in (^{1}, ^{2}) (the upper right number, multiplied by 10^{-8}), and the errors of the grid method (^{7}) with step $h = 0.1$ (the lower left number, multiplied by 10^{-5}) and with step $h = 0.05$ (the lower right number, multiplied by 10^{-5}).

A sharp increase in accuracy was also observed in other numerical examples. In all cases, 24 functions from the potential system were used in the solutions.

Computing Center
Academy of Sciences of the Georgian SSR

Received
27 IX 1968

CITED LITERATURE

1. M. A. Aleksidze, *DAN*, **179**, No. 5 (1968).

2. M. A. Aleksidze, *Journal of Computational Mathematics and Mathematical Physics*, **8**, No. 5 (1968).

3. S. Kaczmarz, G. Steinhaus, *Theory of Orthogonal Series*, Moscow, 1958.

4. N. I. Akhiezer, *Lectures on the Theory of Approximation*, Moscow, 1947.

5. V. N. Faddeeva, *Computational Methods of Linear Algebra*, Moscow-Leningrad, 1950.

6. M. A. Aleksidze, N. L. Lekishvili, *Solution of a Plane Dirichlet Problem by the Method of Generalized Fourier Series*, Moscow, 1967.

7. M. A. Aleksidze, K. V. Pertaya, *Universal Programs for Solving by the Finite-Difference Method the Plane Interior Dirichlet Problem for the Laplace and Poisson Equations*, Tbilisi, 1966.

8. V. D. Kupradze, M. A. Aleksidze, *Journal of Computational Mathematics and Mathematical Physics*, **4**, No. 4 (1964).

9. M. A. Aleksidze, N. M. Arveladze et al., *Communications of the Academy of Sciences of the Georgian SSR*, **49**, No. 2 (1968).

Note: Figure translations are in progress. See original paper for figures.