Influence of various components of errors on the results of approximation using orthogonal functions

V P Zhitnikov¹, N M Sherykhalina¹, N I Zhitnikova¹ and R R Muksimova²

¹ Ufa State Aviation Technical University, 450008, 12, K. Marx str., Ufa, Russia;
² Saint Petersburg State University of Civil Aviation, 196210, 38, Pilotov str., Saint Petersburg, Russia

E-mail: zhitnik@mail.ru

Abstract. The possibility of Legendre polynomials application for approximating functions in order to avoid problems with poorly conditioned matrices for using the least squares method is considered. It is shown that the solutions based on the methods of numerical integration have a large error, which does not allow the idea of orthogonalization to be applied to functions given in a table. However, the use of Legendre polynomials as basic functions instead of algebraic ones in the least squares method can significantly improve the conditionality of matrices. In the problem of approximation of periodic functions by finite Fourier sum the numerical integration is also required. But in connection of the Euler-Maclaurin formula, the numerical integration error of periodic functions is essentially less than of polynomial integration, so the problem is solved very accurately and fast.

1. Introduction

It is experimentally shown in [1,2] that the application of the least squares method (LSM) for approximating functions by algebraic polynomials requires solving systems of linear algebraic equations (SLAE), which are poorly conditioned.

The results of the LSM applying for approximation of the linear function by an algebraic polynomial of degree \( n \) \( P_n(x) = \sum_{i=0}^{n} a_i x^i \) for different numbers of points \((x_i, y_i)\) specifying the approximating function are shown in figure 1,a. Here \( m = n + 1 \) and \( m = 5(n+1) \). The ordinate axis is the decimal logarithms of the errors with the sign «−» This value can be called precision, because it expresses the number of exact decimal digits after points.

The logarithms of the approximation error (the maxima of the difference between the approximating polynomial and the approximated function) are denoted as number 1, the identification errors (the maximum difference between the coefficients of the polynomials obtained as a result of the SLAE solving and the exact values) are denoted as number 2. It can be seen that the identification error can be several orders of magnitude greater than the approximation error and for \( n=11 \) it is about 100%. The further increasing \( n \) does not make sense.

The approximation of other types of functions leads to additional errors, and the error shown in figure 1, a can be considered the minimum value.

It has been noted in [2] the paradoxical relationship between the errors of approximation and identification. Although the coefficients of polynomials with a large error \( \Delta \) are used to determine the...
approximation error, the result of their application for approximation has a significantly smaller error. Explanation of this paradox has been found later and it is given in this work.

This ratio of errors can apparently be explained as follows. Poor conditionality of the SLAE \( a=Ab \) is associated with very large value of the conditioning number \( \nu \)

\[
\frac{|\Delta a|}{|a|} = \nu \frac{|\Delta b|}{|b|}.
\]

The geometric representation of this inequality depends on the type of the applied norm. For example, the inequality \( |\Delta a| \leq \varepsilon \) is equivalent to finding the error vector \( \Delta a \) inside a hypersphere of radius \( \varepsilon \) for Euclidean norm applying. When the norm \( |z| = \max_i |z_i| \) is used the vector \( \Delta a \) lies inside a hypercube with side \( 2\varepsilon \). But neither the first nor the second case can explain the above paradox. However, in the field of interval analysis it has been proved that, in fact, the boundaries of the error range for the solution of SLAE are similar to a star in hyperspace [3]. If the rays of this star are thin enough, this means that for large \( n \) there are such combinations of the sought variables (components of the vector \( a \)), with significant changes in which the values of the penalty function

\[
\Phi(a) = \sum_{j=\lambda}^{m} \left( \sum_{i=0}^{n} a_j x_i^j - y_j \right)^2
\]

undergo very small changes. At the imaginary limit \( n \to \infty \), when the matrix becomes degenerate, there is an infinite set of solutions of the SLAE. So this combination can vary infinitely, and the penalty function remains equal to the minimum.

![Figure 1](image1.png)

**Figure 1.** Dependence on \( n \) of the accuracy of approximation (1) and identification (2) using the SLAE solution: (a) – by Gauss method; (b) – by Jordan-Gauss method. Dotted lines \( y = 16.5 - 0.8n \) and \( y = 17 - 1.5n \).

It should be noted that the difference in identification and approximation errors is observed for using the Gauss method for solving SLAE (reduction to a triangular matrix). If we apply the Jordan-Gauss method for calculating the inverse matrix and multiplying it by the vector of free terms, then the approximation error is practically the same as the identification error [1] (figure 1,b). This is also consistent with the idea of a gradual degeneration of the matrix, since it is possible to analytically solve a SLAE with a degenerate matrix by the Gauss method, despite the multiplicity of solutions, but the inverse matrix does not exist in this case.

Earlier, the idea has been proposed to pass over equations systems solving if the functions are approximated not by an algebraic polynomial, but by polynomials of a special form with the property of orthogonality.
2. Approximation by Legendre polynomials

2.1. The main equations

The main idea of overcoming the problems associated with the conditionality of equations systems arising from the approximation by the least squares method is as follows. Let’s consider the segment \([-1,1]\) to which any finite segment can be reduced using a linear transformation. Let’s consider a system of functions \(G_i(x), \ i = 0, 1, 2, \ldots\) possessing the orthogonality property (equality of the scalar product to zero)

\[
\int_{-1}^{1} G_i(x) G_j(x) dx = 0, \quad i \neq j.
\]  

(1)

The Legendre polynomials have this property and can be defined with help of the recurrence relation

\[
G_0 = 1, \quad G_1(x) = x, \quad G_{i+1}(x) = \frac{2i+1}{i+1} x G_i(x) - \frac{i}{i+1} G_{i-1}(x), \quad i \geq 2.
\]  

(2)

It is proved that any function \(f(x)\) satisfying the Lipschitz condition \(|f(x) - f(y)| \leq L|x - y|, \ L > 0\) can be represented as a convergent series

\[
f(x) = \sum_{i=0}^{\infty} c_i G_i(x).
\]  

(3)

The coefficients of the series \(c_i\) can be obtained by integrating

\[
\int_{-1}^{1} f(x) G_j(x) dx = \sum_{i=0}^{\infty} c_i \int_{-1}^{1} G_i(x) G_j(x) dx = c_j \int_{-1}^{1} G_j^2(x) dx = \frac{2}{2j+1} c_j.
\]

Hence, the coefficients of the series are obtained explicitly

\[
c_i = \frac{2i+1}{2} \int_{-1}^{1} f(x) G_i(x) dx.
\]  

(4)

To solve the problem of approximating a function given by a finite set of values \(x_j, y_j, j=0, \ldots, m\), it is necessary to quantize the solution (3), (4). At first glance, the obvious way is to replace the infinite sum by finite sums (3)

\[
f(x) \approx \sum_{i=0}^{n} c_i G_i(x)
\]  

(5)

and integrals (4) using the trapezoidal method

\[
c_i = \frac{2i+1}{2} \int_{-1}^{1} f(x) G_i(x) dx
\]

(6)

It is convenient to use the matrix of the coefficients of the Legendre polynomials to calculate the coefficients of algebraic polynomials. For \(n = 1\) this matrix has the form

\[
G = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.
\]
For $i \geq 1$, the following recurrent expressions are valid, which follow from (2)

$$
\begin{align*}
& g_{i+1,0} = - \frac{i}{i+1} g_{i-1,0}, \\
& g_{i+1,j+1} = 2\frac{i+1}{i+1} g_{i,j} - \frac{i}{i+1} g_{i-1,j+1}, \quad j = 0, \ldots, i.
\end{align*}
$$

(7)

The coefficients of algebraic polynomials are obtained by the scalar product of the elements of the corresponding columns of the matrix $G$ by the vector $(c_i, i = 1, \ldots, n)$.

### 2.2. Computational experiment

The results of formula (6) using for the approximation of functions $f(x) = 1 + x$ and $f(x) = G_2(x) = 1 - x^2$ for $m = 6(n + 1)$ are shown in the figure 2.

Note that the application of approximate methods for integrals calculating leads to an error that increases with an enlargement of the integrating polynomial degree. In this regard, the accuracy of the obtained approximation results does not exceed two significant digits for $n = 1, 2$ and decreases with an increase of the degree $n$ of the approximating Legendre polynomials, despite the convergence of series (3).

Note that the coefficients in the expansion (5) (lines 3) are identified more accurately than the coefficients of algebraic polynomials obtained from expressions (7) (lines 2).

If the values of the approximated function have a significant component of the random error, the using of the integration method of a higher accuracy order is not advisable, and the method of rectangles can be applied

$$
c_i \approx \frac{2i + 1}{2} \sum_{j=0}^{m-1} f(x_j) G_i(x_j).
$$

However, as a result of this modification of the method, the accuracy of the approximation and identification is less than one significant digit.

![Figure 2](image)

**Figure 2.** Dependence on $n$ of the accuracy of approximation (1), identification of the coefficients of the algebraic polynomial (2) and identification of the coefficients of the Legendre polynomial (3): (a) – linear function; (b) – Legendre polynomial of 2nd degree. Dotted lines $y = 1.5 - n/10$, $y = 2.4 - n/2.5$, $y = 1.8 - n/10$ and $y = 2.4 - n/2.5$.

### 2.3. Method modernization

It is possible to slightly modernize the way integrals are calculated. We interpolate the function $f(x)$, given as a table by a polyline $y = a_j + b_j x$, $x \in [x_j, x_{j+1}]$, $j = 0, \ldots, m - 1$, where
\[ b_j = \frac{y_{j+1} - y_j}{x_{j+1} - x_j}, \quad a_j = y_j - b_j x_j. \]

The integration (4) is carried out according to the formula

\[
\frac{1}{n} \int f(x)G_i(x)dx = \sum_{j=0}^{m-1} \int (a_j + b_j x)G_i(x)dx = \sum_{j=0}^{m-1} \left[ a_j \int_{x_j}^{x_{j+1}} G_i(x)dx + b_j \int_{x_j}^{x_{j+1}} xG_i(x)dx \right].
\]

Integrals on the intervals \([x_j, x_{j+1}]\), taking into account (2), are integrals of polynomials and are calculated theoretically exactly. However, a rounding error takes place. To estimate this error the results of approximation of a linear function (curves 1, denoting practically coinciding errors of approximation and identification) are shown in figure 3. When the quadratic function \(y = 1 + x + x^2\) is approximated, the error of the method occurs and depends little on \(n\) (curves 2 and 3). The error of the method decreases (figure 4, b) with an increase in the number of points \(m\), which is associated with a decrease of the grid step.

**Figure 3.** Dependence on \(n\) of the accuracy of approximation and identification of the linear function (1); identification of the coefficients of the algebraic polynomial (2) and the approximation (3) of the function \(y = 1 + x + x^2\): (a) \(-m = 2 (n + 1)\); (b) \(-m = 16 (n + 1)\). Dotted line \(y = 16 - 0.35n\).

**Figure 4.** Dependence on \(k\) \((m = k(n + 1))\) for the function \(y = 1 + x + x^2\): (a) – identification accuracy; (b) – approximation accuracy.
For a more detailed study of the dependence of the method error on the number of given points \( m \), a sequence of calculations is constructed for approximating the function \( y = 1 + x + x^2 \) for \( n=11 \), \( m = k(n+1), k = 1,2,\ldots \) (figure 4).

The results of calculations and their filtering by the method [4,5,6] show the presence of components of the identification error of the 4th, 6th, 8th orders. The components of the approximation error are of the 2nd, 4th, 6th orders. The investigations have shown that similar orders have errors of approximating the function \( y = 1 + x + x^2 + x^3 \).

For non-polynomial functions approximating, for example \( y = \sin x \), the refinement method by increasing the number of points \( m \) leads to results of limited accuracy (figure 5), since the number \( n=16 \) is fixed. The limitation especially affects the identification accuracy (figure 5,a).

As the experimental results show, an increase of \( n \) also does not lead to an increase of the identification accuracy.

**Figure 5.** Dependence of accuracy on \( k \) \( (m = k(n+1)) \) for the function \( y = \sin x \): (a) – identification accuracy; (b) – approximation accuracy.

Interpolation of tabular values by a polynomial of higher degree than 1 is not promising, since the random part of the error leads to a sharp increase of the error.

Thus, since one should not expect an accuracy increase in solving problems that are important for practice, the application of an explicit method for calculating the coefficients of the sum (5) should be considered unacceptable.

3. The least squares method application

3.1. The least squares method modernization

The second way of quantization is to solve the problem by the least squares method

\[
\Phi(c_0,\ldots,c_n) = \sum_{j=k}^{m} \sum_{i=0}^{n} c_i G_j(x_j) - y_j \right)^2 \rightarrow \min. \tag{8}
\]

The problem is reduced to solving the SLAE

\[
\sum_{i=0}^{n} c_i \sum_{j=0}^{m} G_i(x_j) G_k(x_j) = \sum_{j=0}^{m} y_j G_k(x_j), \quad k = 0,\ldots,n, \quad (m \geq n+1). \tag{9}
\]
Note that the matrix of this system is not diagonal and does not possess the property of diagonal dominance despite the orthogonality of the Legendre polynomials \( G_i(x) \). Moreover, as \( n \) increases, the conditionality of the matrix deteriorates significantly.

The results of approximation and identification of algebraic polynomials of the 1st and \( n \)-th degrees for \( m=6(n+1) \) are shown in figure 6. The similar results of approximation and identification of Legendre polynomials of the 2nd and \( n \)-th degrees are shown in figure 7. One can see the significant accuracy increase in comparison with the results of applying the first quantization method (figure 2). However, as \( n \) increases, the accuracy deteriorates.

**Figure 6.** Dependence on \( n \) of the accuracy of approximation (1), identification of the coefficients of the algebraic polynomial (2) and identification of the coefficients of the Legendre polynomial (3): (a) – linear function; (b) – functions \( x^n \). Dotted lines \( y=15.5 – 0.0185n \), \( y=18 – 0.185n \) (1) and \( y=17 – 0.37n \) (2).

**Figure 7.** Dependence on \( n \) of the approximation accuracy (1), identification of the coefficients of the algebraic polynomial (2) and identification of the coefficients of the Legendre polynomial (3): (a) – Legendre polynomial of the 2nd degree; (b) – Legendre polynomial of the \( n \)-th degree. Dotted lines \( y=15.5 – 0.0185n \), \( y=18 – 0.185n \) (1) and \( y=17 – 0.37n \) (2).

Comparing figures 6 a, b and 3 a, b, it can be seen that the dependences of the approximation and identification errors of the coefficients of the sum (5) differ for different approximated functions.

Since the matrix of the equations system is the same for all the considered examples, it follows that the deterioration in accuracy is associated not with the conditionality of the SLAE matrix, but with the
calculation of the sums and the using of recurrence relations, which contribute to the accumulation of errors.

The errors of approximation (the line 1) and identification of the coefficients of the Legendre polynomials (the line 3) differ little, and the error of approximation slightly exceeds the error of identification. This can be explained by the fact that the values of the polynomial are calculated with the help of the obtained values of the coefficients.

Comment: in contrast to the usual least squares method, in which the approximation error turned out to be significantly less than the identification error, this paradox does not take place in this method.

The identification error of the algebraic polynomial coefficients (the line 2), determined by summing the elements of the columns of the matrix $G$, grows much faster than the approximation errors. The experiment shows that the errors of the coefficients with the middle numbers grow more rapidly. This, apparently, is associated with large values of the elements of the matrix $G$. It follows from a comparison of figure 1 a and 6, that the effect of using orthogonal functions is to reduce the rate of growth of rounding errors with increasing $n$. So, the regression coefficient of the dependence of the identification accuracy of the coefficients of algebraic polynomials on $n$ is 4 (figures 6, 7,a) or 2 (figure 7,b) times less than the same coefficient for the usual least squares method. The decrease of the regression coefficient of the approximation accuracy can be more significant (by a factor of 2 - 40).

3.2. Distortion of the original data by random error

Let’s add a random number $\Delta y_j = c_\Delta \xi_j$, $c_\Delta = 10^k$, $k = -4, -10$, $m = 6(n+1)$ to the initial data $y_j$. Here $\xi_j$ is a quasi-random value uniformly distributed over the interval $(-1, 1)$. The system of equations (9) is solved by the Gauss method. The result of this experiment is shown in figure 8, where the number 1 denotes the dependence of the approximation accuracy of a linear function in comparison with the undistorted result, the number 2 is the identification accuracy of the coefficients of an algebraic polynomial, the number 3 is the accuracy of identification of the coefficients of the Legendre polynomial.

The studies have shown that a random error leads to an almost parallel shift of the lines of the graphs of dependences of accuracy on $n$. The dependence of the approximation accuracy on $n$ approaches to a straight line $y = -k + 0.5 - 0.0185n$ ($c_\Delta = 10^k$), the identification accuracy of the coefficients of an algebraic polynomial approaches to a straight line $y = -k + 2 - 0.37n$.
4. Approximation of periodic functions

4.1. The method specialty

In connection with the large error of numerical integration for direct obtaining the expansion coefficients (5), the question arises how things are in the integration of periodic functions, which is widely applied in spectral analysis.

To approximate periodic functions the representation in the form of a finite Fourier sum is used

\[ f(x) \approx \sum_{i=1}^{n} c_i \sin(i \pi x) + \sum_{i=0}^{n} d_i \cos(i \pi x). \]

Due to the orthogonality of the functions \( \sin(i \pi x) \) and \( \cos(j \pi x) \) (see (1)), the coefficients \( c_i \) and \( d_i \) are defined explicitly

\[ c_i = \int_{-1}^{1} f(x) \sin(i \pi x) \, dx, \quad d_i = \int_{-1}^{1} f(x) \cos(i \pi x) \, dx, \quad i = 1, \ldots, n, \quad d_0 = \frac{1}{2} \int_{-1}^{1} f(x) \, dx. \]

Figures 9, 10 show that the error is significantly less than for integrating the Legendre polynomials (figures 2, 3). This is explained by the small error of integrating the functions \( \sin(i \pi x) \) and \( \cos(i \pi x) \) over the period by the method of trapezoids and rectangles. Indeed, the error estimate for the left rectangle method is represented as the sum (the Euler-Maclaurin formula [7])

\[ \frac{1}{h} \sum_{j=0}^{n-1} f_j - \int_{a}^{b} f(x) \, dx = \sum_{k=1}^{\infty} \gamma_k \left( f^{(k-1)}(b) - f^{(k-1)}(a) \right) h^k + O(h^{k+1}), \]

and all the odd-numbered coefficients \( \gamma_k \), except for the first one, are equal to zero.

The coefficients \( \gamma_k \) satisfy the relation

\[ \frac{x}{e^x - 1} = \sum_{i=0}^{\infty} \gamma_i x^i, \]

and can be obtained by the Bernoulli numbers \( B_k = \gamma_k k! \), which can be obtained from the recurrent expression

\[ B_0 = 1, \quad (k - 1)B_k = -1 - C_{k+1}^1 B_1 - C_{k+1}^2 B_2 - \cdots - C_{k+1}^{k-1} B_{k-1} = - \sum_{i=0}^{k-1} C_i^{k-1} B_i. \]

For the method of right rectangles the estimate is obtained in a similar way. For this, we can replace \( x \) by \( -x \) in the integral. In this case, the right rectangles become the left ones. Under the reverse change in (12) only the first term changes sign (since \( a \) and \( b \) are interchanged), since all other nonzero terms contain only derivatives of odd order.

The trapezoid formula is obtained in the form of a half-sum of two approximate values of the integral calculated by the methods of left and right rectangles. Therefore, the error estimate for the trapezium method is obtained from the right-hand side of (13), if we exclude from it the term with \( k=1 \) (since it is the only one that changes sign when the left rectangles are replaced by the right ones). The error estimate of this type for the trapezium method is given in [7].
Note that for periodic functions, all power terms of \( h \) in (13) are equal to zero, and the error of the method determined by the remainder has a non-power (exponential, factorial) character. These are the reasons for the high accuracy.

Thus, we can state that the approximation of periodic functions by a finite Fourier sum is performed more accurately and faster in comparison with the methods of approximation by algebraic polynomials and Legendre polynomials. In this case (in contrast to the problem of approximation by algebraic polynomials), the identification error, as a rule, is less than the approximation error.

4.2. The computational experiment

The results of applying formulas (10), (11) to solve the problem of approximating functions

\[
 f_1(x) = \sum_{i=1}^{n} \sin(i\pi x) \quad \text{and} \quad f_2(x) = \sum_{i=0}^{n} \cos(i\pi x)
\]

are shown in figure 9.

**Figure 9.** Dependence on \( n \) of the accuracy of approximation (1) and identification (2): (a) – the function \( f_1(x) \); (b) – the function \( f_2(x) \). Dotted lines \( y = 16 - 1.61\lg n \), \( y = 16.25 - 1.21\lg n \).

**Figure 10.** Dependence on \( n \) of the accuracy of approximation (1) and identification (2) of the function \( f_3(x) \): (a) – by the method of rectangles; (b) – by the method of rectangles with acceleration. Dotted lines \( y = 16 - 1.61\lg n \) and \( y = 16.25 - 1.21\lg n \).

Figure 10a shows the results of approximating the function

\[
 f_3(x) = \sum_{i=1}^{n} \sin(i\pi x) + \sum_{i=0}^{n} \cos(i\pi x)
\]

with integration by the method of rectangles, which coincide with the results of integration by the
method of trapeziums up to the rounding error due to their periodicity. To reduce the number of calculations of trigonometric functions, recurrence relations are often used in
\[
\sin((i + 1)\pi x) = \sin(i\pi x) \cos(\pi x) + \cos(i\pi x) \sin(\pi x),
\cos((i + 1)\pi x) = \cos(i\pi x) \cos(\pi x) - \sin(i\pi x) \sin(\pi x).
\]

The calculation results by this method are shown in figure 10 b. It is seen that the differences from figure 10 a are insignificant.

4.3. Distortion of the original data by random error

The influence of a random addition \( \Delta y_j = c_\Delta z_j \), \( c_\Delta = 10^k \), \( m = 6(n+1) \) to the initial data \( y_j \) is illustrated in figure 11, where the lines denoted by roman numerals I, II, III correspond to \( k = -2, -6, -10 \). It is seen that in this range of errors, the result is determined not by the number of terms \( n \), but by the level of random error \( -\lg \Delta \approx k+1.5 \).

Due to the smallness of errors associated with integration and rounding, the main contribution to the errors of approximation and identification is made by the error associated with the cut off of the Fourier series (and the random error of the initial data, if it occurs, as in figure 11).

-\( \lg \Delta \)

-\( \lg n \)

(a)

(b)

**Figure 11.** Dependence on \( n \) of precision of approximation and identification in case of random distortion of the initial data: (a) – the function \( f(x) = \sum_{i=1}^{n} \sin(i\pi x) \); (b) – the function \( f(x) = \sum_{i=0}^{n} \cos(i\pi x) \).

The results of approximating a smooth function \( f_1(x) = x(1 - x^2), x \in [-1,1] \) (the 2nd derivative is discontinuous) and a continuous function \( f_2(x) = \begin{cases} 1 + x, x \in [-1,0] \\ 1 - x, x \in [0,1] \end{cases} \) (the symmetric saw, the 1st derivative is discontinuous) are demonstrated in figure 12.

The Fourier series for these functions are
\[
f_1(x) = \sum_{i=1}^{\infty} \frac{12}{(i\pi)^2} (-1)^i \sin(i\pi x),
\]
\[
f_2(x) = \frac{1}{2} + \sum_{i=1}^{\infty} \frac{2}{((2i-1)\pi)^2} \cos(i\pi x).
\]

It should be noted that for these and other functions with the same features, the above-mentioned properties of the Euler-Maclaurin formula of the equality to zero of all power components of the error of numerical methods do not satisfy.
Figure 12. Dependence on $n$ of the accuracy of approximation (1) and identification (2): (a) – the function $f_1(x) = x(1 - x^2)$; (b) – the symmetric sawtooth function. Dotted lines $y = 0.8 + 2\lg n$, $y = 2.7 + 3\lg n$, $y = 0.75 + 1\lg n$, $y = 1.4 + 2\lg n$.

It is easy to see that the approximation error for $f_1(x)$ has the 2nd order of accuracy, the identification error has the 3rd order. For $f_2(x)$ the approximation error has the 1st order of accuracy, the identification error has the 2nd order. Although these dependences have a power-law character, the errors of numerical integration are much smaller than those in the integration of power functions (figures 2,3). Apparently, this can be explained by the alternation of signs of the integrated functions and errors of integration.

These dependencies can be repeatedly filtered (figures 13, 14) with the refinement of the result by several orders.

Figure 13. Results of filtering the approximation error: (a) – the function $f_1(x)$; (b) – the function $f_2(x)$.

Figure 14. Results of filtering the identification error: (a) – the function $f_1(x)$; (b) – the function $f_2(x)$. 
Figure 15 shows the results of the approximation of two discontinuous functions: \( f_3(x) = x, x \in [-1, 1] \) (the asymmetric saw) and \( f_4(x) = \text{sign} x, x \in [-1, 1] \) (the rectangular pulse). The Fourier series for these functions are

\[
f_3(x) = -\sum_{i=1}^{\infty} \frac{2}{i\pi} (-1)^i \sin(i\pi x), \quad f_4(x) = \sum_{i=1}^{\infty} \frac{4}{(2i-1)\pi} \sin((2i-1)\pi x).
\]

It should be noted that for integrating discontinuous functions, one should take into account their features. So the coefficients for \( f_3(x) \) are obtained by integrating the trapezoid method, for \( f_4(x) \) the method of left rectangles was used in the left half of the segment, and the method of right rectangles in the right half of the segment. As seen from figure 15 the accuracy order of the method is the first for approximation and identification errors.

![Figure 15](image1.png)

The results of filtering data obtained by approximation and identification are shown in figures 16, 17. In the results of approximation the one component of the first-order error is detected and removed, the others have an irregular nature and cannot be filtered.

![Figure 16](image2.png)
The identification error retains a multicomponent structure, despite the discontinuities of the approximated functions, which allows increasing the accuracy by many orders of magnitude using filtering.

5. Conclusion
It is shown that the application of approximate methods for calculating integrals leads to an error in calculating the expansion coefficients of the desired function in terms of the system of the Legendre polynomials, which increases with an increase of the degree of the integrable polynomial. In this regard, the use of an explicit method for calculating the coefficients should be considered unacceptable. However, it is possible to use the least squares method using the Legendre polynomials instead of power terms.

In contrast to the usual least squares method, the conditionality of the matrix limits the application of the method for very large $n$ ($n>100$). There are the other reasons for the accumulation of rounding errors (calculation of sums, recurrent expressions with large coefficients), which limit the possible values of $n$ in the range $n\leq 40 - 50$, depending on the type of approximated functions. In particular, this is due to an increase of the identification error of the algebraic polynomials coefficients. The presence of random perturbation of the initial data significantly increases the errors.

It should be noted that values of $n>10$ are rarely used to solve practical problems. Therefore, the limitations associated with round-off error may not be significant. However, in contrast to the usual least squares method, the error values significantly depend on the type of the approximated function. Therefore, a computational experiment should be carried out using specific functions.

The approximation of periodic functions by a finite Fourier sum is performed more accurately and faster in comparison with the methods of approximation by algebraic polynomials and the Legendre polynomials. In this case (in contrast to the problem of approximation by algebraic polynomials), the identification error, as a rule, is less than the approximation error, it is even possible the increase of the accuracy order.

Acknowledgments
This work was supported by the Russian Foundation for Basic Research (Project code 17-07-00356).

References
[1] Muksimova R R, Zhitnikov V P 2016 Approximation and Identification Errors Investigation Proc. Int. Conf. Mathematical and Informational Technologies (MIT 2016) (Vrnachka Banya, Budva, Serbiya) pp 99-100
[2] Zhitnikov V P, Sherykhalina N M, Muksimova R R and Zhitnikova N I 2020 Influence of different components of data error on the result of solving identification and approximation
problems \textit{IOP Conf. Ser.: Mater. Sci. Eng.} \textbf{919} 052016

[3] Shary S P On optimal solution of interval linear equations 1995 \textit{SIAM Journal on Numerical Analysis} \textbf{32}(2) 610-30

[4] Sherykhalina N M 2007 Methods for processing of the numerical experiment results for their accuracy and reliability increasing \textit{Vestnik UGATU} \textbf{9}(2) 127-37

[5] Zhitnikov V P, Sherykhalina N M 2007 Justification of filtration methods of the numerical experiment results \textit{Vestnik UGATU} \textbf{9}(3) 71-9

[6] Zhitnikov V P, Sherykhalina N M and Sokolova A A 2015 Problem of Reliability Justification of Computation Error Estimates \textit{Mediterranean J. Social Sciences} \textbf{6}(2) 65-78

[7] Bakhvalov N S, Zhidkov N P and Kobel’cov G M 2008 \textit{Numerical Methods} (Moscow: BINOM)