LECTURES

LESSON VIII

8. Approximations of Functions

8.1. Introduction

This chapter is devoted to approximations of functions most applied in different areas of sciences and engineering.

Let function $f : [a, b] \to R$ given by set of value pairs (x_j, f_j) $(j = 0, 1, \ldots, m)$ where $f_j \equiv f(x_j)$. Consider the problem of approximation of function f by linear approximation function

$$\Phi(x) \equiv \Phi(x; a_0, \ldots, a_n) = \sum_{i=0}^{n} a_i \phi_i(x),$$

where $m > n$ (for $m = n$ we have interpolation). Proceeding like at interpolation, we get so known overdefined system of equations

$$\sum_{i=0}^{n} a_i \phi_i(x_j) = f_j \qquad (j = 0, 1, \ldots, m), \tag{8.1.1}$$

which in general case does not have solution, i.e. all equations of system (8.1.1) can not be contemporary satisfied. If we define δ_n by

$$\delta_n(x) = f(x) - \sum_{i=0}^{n} a_i \phi_i(x), \tag{8.1.2}$$

it is possible to search for "solution" of system (8.1.1) so that

$$||\delta_n^*|| = \min_{a_i} ||\delta_n||_r, \tag{8.1.3}$$

where

$$||\delta_n||_r = (\sum_{j=0}^{m} |\delta_n(x_j)|^r)^{1/r} \qquad (r \geq 1).$$

The equality (8.1.3) gives the criteria for determination of parameters $a_0, a_1, \ldots, a_n$ in approximation function Φ. The quantity $||\delta_n^*||_r$, which exists always, is called the value of best approximation in l^r. Optimal values of parameters $a_i = \bar{a}_i$ $(i = 0, 1, \ldots, n)$ in sense of (8.1.3) give the best l^r approximation function

$$\Phi(x)^* = \sum_{i=0}^{n} \bar{a}_i \phi_i(x).$$

Most frequently is taken

1. $r = 1$, $||\delta_n||_1 = \sum_{j=0}^{m} |\delta_n(x_j)|$ (best l^r approximation),

2. $r=2,\ ||\delta_n||_2 = (\sum_{j=0}^{m} |\delta_n(x_j)^2)^{1/2}|$ (mean-square approximation),
3. $r=+\infty,\ ||\delta_n||_\infty = \max_{0\le j\le m}|\delta_n(x_j)|$ (Tchebyshev min-max approximation). In a similar way can be considered problem of best approximation of function f in space $L^r(a,b)$. Here we have

$$||\delta_n||_r = \Big(\int_a^b |\delta_n(x)|^r\, dx\Big)^{1/r} \quad (1 \le r < +\infty)$$

and

$$||\delta_n||_\infty = \max_{a\le x\le b} |\delta_n(x)|.$$

By introducing weight function $p:(a,b)\to R^+$ the more general case of mean-square approximations can be considered, where the corresponding norms for discrete and continuous case are given as (see [1], pp. 90-91)

$$||\delta_n||_2 = ||\delta_n||_{2,p} = \Big(\sum_{j=0}^{m} p(x_j)\delta_n(x_j)|^2\Big)^{1/2} \tag{8.1.4}$$

and

$$||\delta_n||_2 = ||\delta_n||_{2,p} = \Big(\int_a^b p(x)\delta_n(x)^2\, dx\Big)^{1/2}, \tag{8.1.5}$$

respectively.

Example 8.1.1. *Function $x \to f(x) = x^{1/3}$ is to approximate with function $x \to \phi(x) = a_0 + a_1 x$ in space*

$$1^0 \ \ L^1(0,1), \qquad 2^0 \ \ L^2(0,1), \qquad 3^0 \ \ L^\infty(0,1).$$

Here we have $\delta_1(x) = x^{1/3} - a_0 - a_1 x \quad (0 \le x \le 1)$. (see [1], pp. 91-93).

1^0 We get the best $L^1(0,1)$ approximation by minimization of norm

$$||\delta_1||_1 = \int_0^1 |x^{1/3} - a_0 - a_1 x| dx.$$

Having $\dfrac{\partial\delta_1}{\delta a_0} = -1$, and $\dfrac{\partial\delta_1}{\delta a_1} = -x$, the optimal values of parameters a_0 and a_1 are to be determined from the system of equations

$$\begin{aligned} &\int_0^1 sgn\delta_1(x)\, dx = 0,\\ &\int_0^1 x\, sgn\delta_1(x)\, dx = 0. \end{aligned} \tag{8.1.6}$$

Having in mind that δ_1 changes sign on segment $[0,1]$ in points x_1 and x_2 (see fig. 8.1.1), system of equations 8.1.6 reduces to system

$$x_2 - x_1 = \frac{1}{2}, \qquad x_2^2 - x_1^2 = \frac{1}{2},$$

wherefrom it follows $x_1 = \frac{1}{4}$ and $x_2 = \frac{3}{4}$.

Thus, determining the best $L^1(0,1)$ approximation reduces to interpolation, i.e. determining of interpolation polynomial Φ^* which satisfies the conditions

$$\Phi^*(1/4) = f(1/4) = \sqrt[3]{\frac{1}{4}}, \quad \Phi^*(3/4) = f(3/4) = \sqrt[3]{\frac{3}{4}},$$

i.e.

$$\begin{aligned}\Phi^*(x) &= \frac{2}{3\sqrt[3]{4}}(\sqrt[3]{3}-1)x + \frac{1}{2\sqrt[3]{4}}(3-\sqrt[3]{3})\\ &\cong 0.55720x + 0.49066.\end{aligned}$$

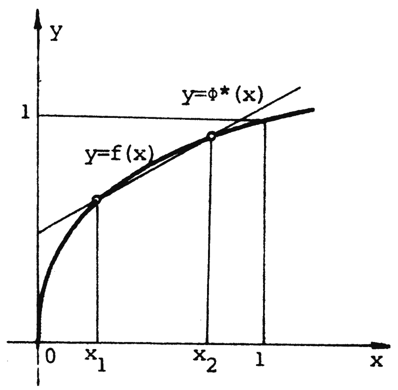

Fig. 8.1.1

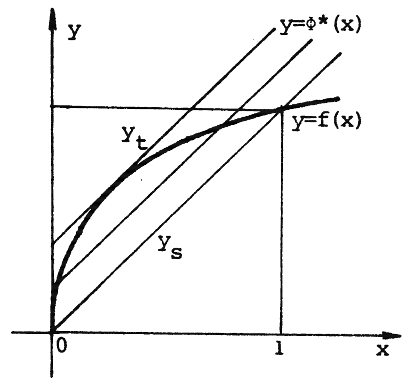

Fig. 8.1.2

2^0 Let

$$I(a_0,a_1) = ||\delta_1||_2^2 = \int_0^1 (x^{1/3} - a_0 - a_1 x)^2 dx.$$

From the conditions

$$\frac{\partial I}{\partial a_0} = -2\int_0^1 (x^{1/3} - a_0 - a_1 x)dx = 0,$$

$$\frac{\partial I}{\partial a_1} = -2\int_0^1 x(x^{1/3} - a_0 - a_1 x)dx = 0,$$

it follows

$$a_0 + \frac{1}{2}a_1 = \frac{3}{4}, \quad \frac{1}{2}a_0 + \frac{1}{3}a_1 = \frac{3}{7},$$

i.e. $a_0 = \overline{a_0} = \frac{3}{7}$, $a_1 = \overline{a_1} = \frac{9}{14}$, so that the best mean-square approximation is given with

$$\Phi^*(x) = \frac{3}{7} + \frac{9}{14}x \cong 0.42857x + 0.64286x.$$

3^0 For determining of min-max approximation we will use the following simple geometrical procedure. Through the end-points of curve $y = f(x) = x^{1/3}$ $(0 \geq x \geq 1)$ we will draw the secant, and then tangent on the curve which is parallel to this secant (see Fig. 8.1.2). The corresponding equations for those straight lines are

$$y = y_{sec} = x \quad \text{and} \qquad y = y_{tan} = x + \frac{2\sqrt{3}}{9},$$

so that the best min-max approximation is

$$\Phi^*(x) = \frac{1}{2}(y_{sec} + y_{tan}) = x + \frac{\sqrt{3}}{9} \cong x + 0.19245,$$

whereby the the value of best approximation is $||\delta_1^*||_\infty = \frac{\sqrt{3}}{9}$.

8.2. Mean-square approximation

Here, we will consider the problem of best approximation of function $f : [a, b] \to R$ using linear approximation function

$$\Phi(x) = \sum_{i=0}^{n} a_i \Phi_i(x),$$

where $\{\Phi_i\}$ is system of linear independent functions from the space $L^2(a, b)$, with scalar product introduced by

$$(f, g) = \int_a^b p(x)f(x)g(x)dx \quad (f, g \in L^2(a, b)),$$

where $p : (a, b) \to R^+$ is given weight function.

From the previous section we can conclude that for the best mean-square approximation for f it is necessary to minimize the norm (8.1.5) by parameters $a_i \quad (i = 0, 1, \ldots, n)$.

If we put $I(a_0, a_1, \ldots, a_n) = ||\delta_n||^2 = (\delta_n, \delta_n)$, then from

$$\frac{\partial I}{\partial a_j} = 2\int_a^b p(x)(f(x) - \sum_{i=0}^{n} a_i\Phi_i(x))(-\Phi_j(x))dx = 0 \quad (j = 0, 1, \ldots, n)$$

it follows system of equations for determination of approximation parameters

$$\text{(8.2.1)} \qquad \sum_{i=0}^{n}(\Phi_i, \Phi_j)a_i = (f, \Phi_j) \quad (j = 0, 1, \ldots, n).$$

This system can be represented in matrix form as

$$\begin{bmatrix} (\Phi_0, \Phi_0) & (\Phi_1, \Phi_0) & \ldots & (\Phi_n, \Phi_0) \\ (\Phi_0, \Phi_1) & (\Phi_1, \Phi_1) & & (\Phi_n, \Phi_1) \\ \vdots & & & \\ (\Phi_0, \Phi_n) & (\Phi_1, \Phi_n) & & (\Phi_n, \Phi_n) \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} (f, \Phi_0) \\ (f, \Phi_1) \\ \vdots \\ (f, \Phi_n) \end{bmatrix},$$

Matrix of this system is known as Gram's matrix. It can be shown that this matrix is regular if system of functions $\{\Phi_i\}$ is linearly independent, having unique solution of given approximation problem.

System of equations (8.2.1) can be simple solved if the system of functions $\{\Phi\}$ is orthogonal. Namely, all off-diagonal elements of matrix of system are equal to zero, i.e. matrix is diagonal one, having as solutions

$$\text{(8.2.2)} \qquad a_i = \overline{a_i} = \frac{(f, \Phi_i)}{(\Phi_i, \Phi_i)} \quad (i = 0, 1, \ldots, n).$$

It can be shown that by taking in the given way chosen parameters $a_i \quad (i = 0, 1, \ldots, n)$ the function I reaches its minimal value. Namely, because

$$\frac{\partial^2 I}{\partial a_j \partial a_k} = 2(\Phi_k, \Phi_j) = 2||\Phi_k||^2\delta_{kj},$$

where δ_{kj} is Cronecker delta, we have

$$d^2 I = 2\sum_{i=0}^{n} ||\Phi_k||^2 da_k^2 > 0.$$

Thus, the best mean-square approximation of function f in subspace $X_n = L(\Phi_0, \Phi_1, \ldots, \Phi_n)$, where $\{\Phi_i\}$ is orthogonal system of functions, is given as

$$\Phi^*(x) = \sum_{i=0}^{n} \frac{(f, \Phi_i)}{||\Phi_i||^2} \Phi_i(x). \tag{8.2.3}$$

A very important class of mean-square approximations is approximation by algebraic polynomials. In this case, the orthogonal basis of semi-space X_n is constructed by Gramm-Schmidt orthogonalisation procedure, starting, for example, from natural basis $\{1, x, x^2, \ldots, x^n\}$ (see Chapter IV), or general methods for orthogonalisation.

Example 8.2.1. *For function $x \to f(x) = |x|$ on segment $[-1, 1]$ determine in the set of polynomials not greater degree than three, best mean-square approximation, with weight function $x \to p(x) = (1-x^2)^{3/2}$.*

Let us compute integral

$$N_k = \int_{-1}^{1} x^{2k}(1-x^2)^{3/2}\, dx \quad (k \in N_0)$$

needed for further considerations (see [4], pp. 92-93). By partial integration over the integral

$$N_{k-1} - N_k = \int_{-1}^{+1} x^{2(k-1)}(1-x^2)^{5/2}\, dx \qquad (k \in N),$$

we get $N_{k-1} - N_k = \dfrac{5}{2k-1} N_k$, i.e. $N_k = 3\pi \dfrac{2k-1}{2k+4} N_{k-1}, \quad (k \in N)$ so that, with $N_0 = \dfrac{3\pi}{8}$ we have $N_k = 3\pi \dfrac{(2k-1)!!}{(2k+4)!!} \quad (k \in N)$. Starting from natural basis $\{1, x, x^2, \ldots\}$, using Gramm-Schmidt orthogonalisation, we get subsequently

$$\begin{aligned}
\Phi_0(x) &= 1 \\
\Phi_1(x) &= x - \frac{(x, \Phi_0)}{(\Phi_0, \Phi_0)} \Phi_0(x) = x, \\
\Phi_2(x) &= x^2 - N_1 N_0^{-1} = x^2 - \frac{1}{6}, \\
\Phi_3(x) &= x^3 - N_2 N_1^{-1} x = x^3 - \frac{3}{8} x,
\end{aligned}$$

and corresponding norms

$$||\Phi_0|| = \sqrt{\frac{3\pi}{8}}, \ ||\Phi_1|| = \frac{\sqrt{\pi}}{4}, \ ||\Phi_2|| = \frac{1}{8}\sqrt{\frac{5\pi}{5}}, \ ||\Phi_3|| = \frac{\sqrt{3\pi}}{32}.$$

Because of

$$(f, \Phi_0) = \frac{2}{5}, \ (f, \Phi_1) = 0, \ (f, \Phi_2) = \frac{1}{21}, \ (f, \Phi_3) = 0,$$

using (8.2.2) we get

$$a_0 = \frac{16}{15\pi}, \ a_1 = 0, \ a_2 = \frac{128}{35\pi}, \ a_3 = 0,$$

having, finally approximation in the form

$$\Phi^*(x) = \frac{16}{15\pi} + \frac{128}{35\pi}(x^2 - \frac{1}{6}) \cong 0.14551309 + 1.1641047x^2.$$

This function is, in addition, best approximation in the set of polynomials of degree not greater than two.

Some further very valuable considerations can be found in [1], pp. 96-99.

8.3. Mean-square approximation with boundaries

Mean-square approximations can be applied to different class of functions and weight functions. Taking Gegenbauer weight function $x \to p(x) = (1-x^2)^{\lambda-1/2}$ on $[-1,1]$ for real functions which belong to some class of functions like

$$FE = \{f|\ f(-x) = f(x), f(1) = 0, f \in L^2[-1,1]\}$$

and

$$FO = \{f|\ f(-x) = -f(x), f(1) = 0, f \in L^2[-1,1]\}$$

where to the approximation functions are intruded such limitations so that they belong to the same class of functions. Such approximations are often demanded in practice. Note that FE denotes class of even functions and FO class of odd functions with zero values in points $x = 1$ and $x = -1$.

It is possible to consider more general case of approximation using other weight functions and other boundaries. In this case we take for weight function $p(x)$ Gegenbauer function and boundaries over approximation function such that belongs to the same class of function like function f.

Thus, scalar product is

$$(8.3.1) \qquad (f, g) = \int_{-1}^{1} p(x) f(x) g(x)\, dx \qquad (p(x) = (1-x^2)^{\lambda-1/2},\ \lambda > -1/2)$$

Let further $\mathcal{P}_m$ be set of all algebraic polynomials of degree not grater than m and belong to set FE if m even, and to FO if m odd. For function $f \in FE$ (or FO) we will determine mean-square approximation in class $\mathcal{P}_{2n}$ (or $\mathcal{P}_{2n+1}$) in regard to norm involved by scalar product (8.3.1). So we have that approximations Φ_{2n} and Φ_{2n+1} are solutions of two minimization problems, respectively:

$$(8.3.2) \qquad \min_{\Phi \in \mathcal{P}_{2n}} ||f - \Phi||, \qquad \text{when } f \in FE,$$

and

$$(8.3.3) \qquad \min_{\Phi \in \mathcal{P}_{2n+1}} ||f - \Phi||, \qquad \text{when } f \in FO.$$

In general case, when f is neither even nor odd, but satisfies condition $f(-1) = f(1) = 0$, mean-square approximation ψ_m (in class of polynomials of degree not greater than m), which satisfies the conditions $\psi_m(-1) = \psi_m(1) = 0$ is simple to be obtained as

$$\psi_m(x) = \Phi_{2n}(x) + \Phi_{2n+1}(x),\ \text{when } m = 2n+1,$$
$$\psi_m(x) = \Phi_{2n}(x) + \Phi_{2n-1}(x),\ \text{when } m = 2n,$$

where Φ_{2n} and $\Phi_{2n\pm1}$ are solutions of problems (8.3.2) and (8.3.3), what follows from representation

$$f(x) = \frac{1}{2}(f(x) + f(-x)) + \frac{1}{2}(f(x) - f(-x)).$$

Some advanced results are given by Milovanović and Wrigge in ([6] and [1], pp. 101-105).

8.4. Economization of power series

As previously mentioned, for function evaluation are often used polynomial developments. For example, if function f on segment $[-1,1]$ has development

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots,$$

then for evaluation of function value on segment $[-1, 1]$ can be used the polynomial

$$P_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n. \tag{8.4.1}$$

The procedure of power series economization, arising from Lanczos, consists from lowering of degree of polynomial (8.4.1) with slightly error increasing, and proceeds rather simple, by using orthogonal polynomials. Most frequently are used Chebyshev and Legendre polynomials.

Consider economization by using Chebyshev polynomials T_k $(k = 0, 1, \ldots)$. Having

$$\begin{aligned}
&T_0(x) = 1,\ T_1(x) = x,\ T_2(x) = 2x^2 - 1,\ T_3(x) = 4x^3 - 3x,\\
&T_4(x) = 8x^4 - 8x^2 + 1,\ T_5(x) = 16x^5 - 20x^3 + 5x,\\
&T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1,\ T_7(x) = 64x^7 - 112x^5 + 56x^3 - 7x,
\end{aligned}$$

etc., it is possible to express algebraic degrees x^k $(k = 0, 1, \ldots)$ using Chebyshev basis, in the following way:

$$\begin{aligned}
&1 = T_0,\ x = T_1,\ x^2 = \frac{1}{2}(T_0 + T_2),\ x^3 = \frac{1}{4}(3T_1 + T_3),\\
&x^4 = \frac{1}{8}(3T_0 + 4T_2 + T_4),\ x^5 = \frac{1}{16}(10T_1 + 5T_3 + T_5),\\
&x^6 = \frac{1}{32}(10T_0 + 15T_2 + 6T_4 + T_6),\ x^7 = \frac{1}{64}(35T_1 + 21T_3 + 7T_5 + T_7),
\end{aligned}$$

etc. In general case, it holds

$$x^k = \frac{1}{2^{k-1}} \sum_{i=0}^{[k/2]} \frac{\binom{k}{i}}{1 + \delta_{k,2i}} T_{k-2i}(x).$$

Using these formulas, the polynomial (8.4.1) can be presented in the form

$$P_n(x) = c_0 T_0(x) + c_1 T_1(x) + \cdots + c_n T_n(x). \tag{8.4.2}$$

Denote with $\mathcal{P}_m$ set of all algebraic polynomials of degree not greater than m. Taking first $m + 1$ $(m < n)$ members in development (8.4.2), we get the polynomial

$$Q_m(x) = c_0 T_0(x) + c_1 T_1(x) + \cdots + c_m T_m(x), \tag{8.4.3}$$

which is the approximation of P_n in set $\mathcal{P}_m$.

In regard to fact that Chebyshev polynomials satisfy inequality $|T_k(x) \le 1|$ $(-1 \le x \le 1)$, for approximation error it holds

$$|P_n(x) - Q_m(x)| \le |c_{m+1}| + \cdots + |c_n| \qquad (-1 \le x \le 1).$$

The given approximation procedure is called Lanczos economization (it is suggested to students to write a short-line software for this procedure using `Mathematica`). The following theorem (without proof) explains the kind of approximation.

Theorem 8.4.1. *The polynomial Q_m, given by (8.4.3) represents in set $\mathcal{P}_m$ the best mean-square approximation with Tchebyshev weight function $p(x) = (1 - x^2)^{-1/2}$ for polynomial P_n on segment $[-1, 1]$.*

Example 8.4.1. *Using economization approximate $x \to x^6$ $(|x| \le 1)$ by polynomial of degree not greater than two.*

Having $x^6 = \frac{1}{32}(10T_0 + 15T_2 + 6T_4 + T_6)$, by cutting off the last two members we get the polynomial $Q_2(x) = \frac{10}{32} + \frac{15}{32}(2x^2 - 1) = \frac{15}{16}x^2 - \frac{5}{32}$, whereby the absolute error is lesser than $\frac{7}{32}$, i.e. it holds $|x^6 - Q_2(x)| \le \frac{7}{32}$.

By development using Legendre polynomials (mean-square approximation with weight $p(x) = 1$), the following approximation is obtained.

$$x^6 \cong \frac{1}{7}P_0(x) + \frac{10}{21}P_2(x) = \frac{5}{7}x^2 - \frac{2}{21} \qquad (-1 \le x \le 1)$$

with absolute error not greater than $8/21$. Note that this error is greater, i.e. $8/21 > 7/32$.

8.5. Discrete mean-square approximation

In previous sections we considered problem of best approximation of function in space $L^2(a, b)$. Now we will consider a particular case, mentioned in introductory section. Namely, let function $f : [a, b] \to R$ be given on set of pairs of values $\{(x_j, f_j)\}_{j=0,1,\ldots,m}$, where $f_j \equiv f(x_j)$. We will consider the problem of best approximation of given function by linear approximation function

$$\Phi(x) = \sum_{i=0}^{n} a_i \Phi_i(x) \quad (n < m) \tag{8.5.1}$$

in sense of minimization of norm (8.1.4), where $p : [a, b] \to R_+$ is given weight function and δ_n defined by (8.1.2). By involving the matrix notation

$$\mathbf{X} = \begin{bmatrix} \Phi_0(x_0) & \Phi_1(x_0) & \ldots & \Phi_n(x_0) \\ \Phi_0(x_1) & \Phi_1(x_1) & \ldots & \Phi_n(x_1) \\ \vdots & & & \\ \Phi_0(x_m) & \Phi_1(x_m) & \ldots & \Phi_n(x_m) \end{bmatrix}, \ \vec{f} = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{bmatrix}, \ \vec{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix},$$

$$\mathbf{P} = diag(p(x_0), p(x_1), \ldots, p(x_m)), \quad \vec{v} = \vec{f} - \mathbf{X}\vec{a},$$

square of norm, defined by (8.1.4), can be represented as

$$F = ||\delta_n||^2 = ||\delta_n||_2^2 = \sum_{j=0}^{m} p(x_j)\delta_n(x_j)^2 = \vec{v}^T \mathbf{P} \vec{v}. \tag{8.5.2}$$

For determination of best discrete mean-square approximation (8.5.1)it is necessary to minimize F, given by (8.5.2). Thus, based on

$$\frac{\partial F}{\partial a_i} = 2 \sum_{j=0}^{m} p(x_j)\delta_n(x_j) \frac{\partial \delta_n(x_j)}{\partial a_i} = 0 \qquad (i = 0, 1, \ldots, n)$$

we get normal system of equations

$$\sum_{j=0}^{m} p(x_j)\delta_n(x_j)\Phi_i(x_j) = 0 \qquad (i = 0, 1, \ldots, n) \tag{8.5.3}$$

for determination of parameters a_i, $(i = 0, 1, \ldots, n)$. The last system of equations can be given in matrix form

$$\vec{v}^T \mathbf{P} \vec{v} = \vec{0},$$

i.e.

$$\mathbf{X}^T \mathbf{P} \mathbf{X} \vec{a} = \mathbf{X}^T \mathbf{P} \vec{f}. \tag{8.5.4}$$

Note that normal system of equations (8.5.3), i.e (8.5.4) is obtained from overdefined system of equations (8.1.1), given in matrix form as

$$\mathbf{X}\vec{a} = \vec{f},$$

by simple multiplication by matrix $\mathbf{X}^T\mathbf{P}$ from the left side.

Diagonal matrix $\mathbf{P}$, which is called weight matrix, is of meaning so that larger weights $p_j \equiv p(x_j)$ are assigned to the values of function f_j with greater accuracy. This is of importance when approximating experimental data, which are obtained during measures by different accuracy. For example, for measurements realized with different dispersions, which relations are known, the weights p_j are chosen as inverse of dispersions, i.e, such that

$$p_0 : p_1 : \cdots : p_m = \frac{1}{\sigma_0^2} : \frac{1}{\sigma_1^2} : \cdots : \frac{1}{\sigma_m^2}.$$

When the measurements are realized with same accuracy, but with different numbers of measurements, i.e. for every value of argument x_j are proceeded m_j measurements, and for f_j taken arithmetic means of obtained results in series of measurements, then for weights are taken numbers measurements in series, i.e. $p_j = m_j$ $(j = 0, 1, \ldots, m)$. Nevertheless, usually are the weights equal, i.e. $\mathbf{P}$ is unit matrix of order $m+1$. In this case, (8.5.4) reduces to

$$\mathbf{X}^T\mathbf{X}\vec{a} = \mathbf{X}^T\vec{f}. \tag{8.5.5}$$

Vector of coefficients $\vec{a}$ is determined from 8.5.4 or 8.5.5. From 8.5.5 it follows

$$\vec{a} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\vec{f}.$$

In case when the system of basic functions is chosen so that $\Phi_i(x) = x^i$ $(i = 0, 1, \ldots, n)$ we have

$$\mathbf{X} = \begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & & x_1^n \\ \vdots & & & & \\ 1 & x_m & x_m^2 & & x_m^n \end{bmatrix}.$$

The method considered is often called least-square method. Interesting case is when $n = 1$, i.e. when the approximation function is of form $\Phi(x) = a_0 + a_1 x$. Then the system (8.5.4) becomes

$$\begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix},$$

where

$$s_{11} = \sum_{j=0}^{m} p_j, \;\; s_{12} = s_{21} = \sum_{j=0}^{m} p_j x_j, \;\; s_{22} = \sum_{j=0}^{m} p_j x_j^2,$$
$$b_0 = \sum_{j=0}^{m} p_j f_j, \;\; b_1 = \sum_{j=0}^{m} p_j x_j f_j.$$

The asked approximation parameters are

$$a_0 = \frac{1}{D}(s_{22}b_0 - s_{12}b_1), \;\; a_1 = \frac{1}{D}(s_{11}b_1 - s_{21}b_0),$$

where $D = s_{11}s_{22} - s_{12}^2$.

Example 8.5.1. *Find parameters a_0 and a_1 in approximation function $\Phi(x) = a_0 + a_1 x$ using least-square method, for function given in tabular form, as a set of values pairs*

$$\{(1.1, 2.5),\; (1.9, 3.2),\; (4.2, 4.5),\; (6.1, 6.0)\}.$$

For weight matrix $\mathbf{P}$ we can take unit matrix. The previously given formulas can be directly applied, but we can start from overdefined system of equations

$$\begin{bmatrix} 1 & 1.1 \\ 1 & 1.9 \\ 1 & 4.2 \\ 1 & 6.1 \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 2.5 \\ 3.2 \\ 4.5 \\ 6.0 \end{bmatrix},$$

By multiplying with matrix $\mathbf{X}^T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1.1 & 1.9 & 4.2 & 6.1 \end{bmatrix}$ from the left side, we get the normal system of equations

$$\begin{bmatrix} 4 & 13.3 \\ 13.3 & 59.67 \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 16.2 \\ 64.33 \end{bmatrix},$$

wherefrom it follows

$$\begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \frac{1}{61.79} \begin{bmatrix} 59.67 & -13.3 \\ -13.3 & 4. \end{bmatrix} \cdot \begin{bmatrix} 16.2 \\ 64.33 \end{bmatrix} = \begin{bmatrix} 1.7974591 \\ 0.6774559 \end{bmatrix},$$

Thus, we have $\Phi(x) = 1.7974591 + 0.6774559x$.

In case when $n > 1$, least-square method becomes complicated because the obtained system of linear equations is more difficult to be solved. This system could be simple solved if the system matrix reduces to diagonal matrix, what happens when system $\{\Phi_k\}$ is orthogonal system of polynomials, meaning that all off-diagonal matrix members are equal to zero. These orthogonal polynomials are discrete (see [4], pp. 154-159). Thus, one shell take $\Phi_k(x) = Q_k^{(N)}(x)$ $(k = 0, 1, \ldots, N)$ where $N - 1 = m$ and scalar product defined as

$$(f, g) = [f, g]_N = \sum_{i=0}^{N-1} p_i f(x_i) g(x_i).$$

As already known, series of orthogonal polynomials can be obtained by Sieltjes procedure, i.e. it is possible to determine coefficients $\beta_k^{(N)}$ and $\gamma_k^{(N)}$ in recurrence relation (2.5.15) (see [4], p. 155).

It is suggested to students to write code for very usable Stieltjes procedure for continuous and discrete orthogonal polynomials, using system `Mathematica` (hint: see [8]).

Similar to continuous mean-square approximation, we get the discrete one in form

$$\Phi(x) = \sum_{i=0}^{n} a_i Q_i^{(N)}(x),$$

where approximation parameters are given as

$$a_k = \frac{[f, Q_k^{(N)}]_N}{||Q_k^{(N)}||^2}, \qquad (k = 0, 1, \ldots, n). \tag{8.5.6}$$

In case when points x_i are equidistant (taking, without generality lessening, $x_i = i$ $(i = 0, 1, \ldots, N - 1)$) and with equal weights (for example, $p_i = 1/N$), we have the case of discrete Chebyshev polynomials, with explicit known coefficients of three-term recurrence relation (see [4], pp. 156-157). Mean-square approximations are in this case simple to obtain.

In many areas of science and technology, dealing with experimental data, we have often of parameter determination in so known empirical formulas which express functional relation between two or more variables. For example, when functional relation given as

$$y = f(x; a_0, a_1, \ldots, a_n),$$

where a_i $(i = 0, 1, \ldots, n)$ are parameters which are to be determined using the following tabulated data obtained by measurement.

i	0	1	$\ldots$	m
x_i	x_0	x_1	$\ldots$	x_m
y_i	y_0	y_1	$\ldots$	y_m

The measured data contain accidental errors of measurements, i.e. "noise" in experiment. Determination of parameters a_i $(i = 0, 1, \ldots, n)$ is, from the point of theory of approximation, possible only if $m \geq n$. In case of $m = n$, we have interpolation, which is, in general case nonlinear, what depends on function shape. In order to eliminate "noise" in data, and obtain greater exactness and reliability, the number of measurements should be large enough. Then, the most used method for determination of parameters is least-square method, i.e. minimization of variable F defined by

$$F = \sum_{j=0}^{m} (y_j - f(x_j; a_0, a_1, \ldots, a_n))^2, \tag{8.5.7}$$

or using

$$F = \sum_{j=0}^{m} p_j \, (y_j - f(x_j; a_0, a_1, \ldots, a_n))^2,$$

where are included weights p_j. If functional relation between several variables is given as

$$z = f(x, y; a_0, a_1, \ldots, a_n)$$

for determination of approximation parameters we have to minimize

$$F = \sum_{j=0}^{m} p_j \, (z_j - f(x_j, y_j; a_0, a_1, \ldots, a_n))^2.$$

If f is linear approximation function (in parameters $a_0, a_1, \ldots, a_n$), i.e. of form (8.5.1), the problem is to be solved in previously explained way. Nevertheless, if f is nonlinear approximation function, then the corresponding normal system of equation

$$\frac{\partial F}{\partial a_i} = 0 \qquad (i = 0, 1, \ldots, n) \tag{8.5.8}$$

nonlinear. For solving of this system can be used some method for solution of system of nonlinear equations, like Newton-Kantorovich method, whereby is this procedure rather complicated. In order to solve problem in easier way, there are some simplified methods of transformation of such problems to linear approximation method. Namely, by introducing some substitutions, like

$$X = g(x), \quad Y = h(y) \tag{8.5.9}$$

nonlinear problem reduces to linear one.

For example, let $y = f(x; a_0, a_1) = a_0 e^{a_1 x}$.Then, by logaritmization and substitution

$$X = x, \quad Y = \log y, \quad b_0 = \log a_0, \quad b_1 = a_1,$$

the problem is reduced to linear one, because $Y = b_0 + b_1 X$. Thus, by minimization of

$$G = G(b_0, b_1) = \sum_{j=0}^{m} (Y_j - b_0 - b_1 X_j)^2, \tag{8.5.10}$$

where $X_j = x_j$ and $Y_j = \log y_j$ $(j = 0, 1, \ldots, n)$, we determine the parameters b_0 and b_1, and then

$$a_0 = e^{b_0} \text{ and } a_1 = b_1.$$

Nevertheless, this procedure does not give the same result like the minimization of function

$$F = F(a_0, a_0) = \sum_{j=0}^{m} (y_j - a_0 e^{a_1 x_j})^2.$$

Moreover, the obtained results can significantly deviate, because the problem we are solving is different from stated one, having in mind transformation we have done ($Y = \log y$). But, for many practical engineering problems the parameters obtained in this way are satisfactory.

We will specify some typical functional dependencies with possible transformation of variables.

1^0 $y = a_0 x^{a_1},\ X = \log x,\ Y = \log y,\ b_0 = \log a_0,\ b_1 = a_1;$

2^0 $y = a_0 a_1^x,\ X = x,\ Y = \log y,\ b_0 = \log a_0,\ b_1 = \log a_1;$

3^0 $y = a_0 + \dfrac{a_1}{x},\ X = \dfrac{1}{x},\ Y = y,\ b_0 = a_0,\ b_1 = a_1;$

4^0 $y = a_0 + \dfrac{a_1}{x},\ X = x,\ Y = xy,\ b_0 = a_1,\ b_1 = a_0;$

p

5^0 $y = \dfrac{1}{a_0 + a_1 x},\ X = x,\ Y = \dfrac{1}{y},\ b_0 = a_0,\ b_1 = a_1;$

6^0 $y = \dfrac{x}{a_0 + a_1 x},\ X = \dfrac{1}{x},\ Y = \dfrac{1}{y},\ b_0 = a_1,\ b_1 = a_0;$

7^0 $y = \dfrac{x}{a_0 + a_1 x},\ X = x,\ Y = \dfrac{x}{y},\ b_0 = a_0,\ b_1 = a_1;$

8^0 $y = \dfrac{1}{a_0 + a_1 e^{-x}},\ X = e^{-x},\ Y = \dfrac{1}{y},\ b_0 = a_0,\ b_1 = a_1;$

9^0 $y = a_0 + a_1 \log x,\ X = \log x,\ Y = y,\ b_0 = a_0,\ b_1 = a_1.$

Example 8.5.3. *Result of measurements of values x and y are given in following tabular form.*

i	0	1	2	3	4	5
x_i	4.48	4.98	5.60	6.11	6.62	7.42
y_i	4.15	1.95	1.31	1.03	0.74	0.63

If $y = \dfrac{1}{a_0 + a_1 x}$, *reduce to linear problem and approximate using least-square method.*

By involving $X = x,\ Y = 1/y$ and using least-square method we get approximation function $\Phi(X) \cong 0.468X - 1.843$, wherefrom it follows $y \cong \dfrac{1}{0.468x - 1.843}$.

From the previous one can conclude that, depending on f the convenient replacements (3.5.9) should be chosen so that they enable reducing of

$$y = f(x; a_0, a_1, \ldots, a_n)$$

to linear form of, for example, polynomial type

$$Y = b_0 + b_1 X + \ldots + b_n X^n. \tag{8.5.11}$$

It is clear that functions g and h must have their inverse functions, so that (8.5.11) is, in fact, equivalent to

$$h^{-1}(Y) = f(g^{-1}(X); a_0, a_1, \ldots, a_n),$$

whereby parameters b_i depend on parameters a_i in rather simple way.

8.6. Chebyshev min-max approximation

In this section will be given basics of $\min - \max$ approximation of function $f \in C[a, b]$ by algebraic polynomials. Most of results can be translated to more general types of approximating functions.

Let denote with $\mathcal{P}_n$ set of all algebraic polynomials of degree not greater than n. We have the problem of determination of polynomial $P_n = P_n^*(in \mathcal{P}_n)$ which minimizes norm $||f - P_n||_\infty$. Thus, the minimization problem of form

$$E_n(f) = \min_{P_n \in \mathcal{P}_n)} (\max_{a \le x \le b} |f(x) - P_n(x)|) = \max_{a \le x \le b} |f(x) - P_n^*(x)|,$$

where $E_n(f)$ is the value of best approximation, is to be solved.

Algebraic polynomials are very good approximation elements on finite segment $[a, b]$ and therefor often used at $\min - \max$ approximations, and at approximations at all. Weierstrass' theorem defines existence of polynomials of large enough degree, which has arbitrary small deviation from continuous function on $[a, b]$. The following theorem, nevertheless, gives criterion for statement that given polynomial is the best $\min - \max$ approximation of given continuous function on $[a, b]$ in class of algebraic polynomials $\mathcal{P}_n$.

Theorem 8.6.1. *The polynomial P_n^* is the best* $\min - \max$ *approximation of function $f \in [a, b]$ in set $\mathcal{P}_n$ if and only if on segment $[a, b]$ exist $n + 2$ points $x_0, x_1, \ldots, x_{n+1}$ $(x_0 < x_1 < \cdots < x_{n+1})$ such that*

$$\delta_n^*(x_k) = -\delta_n^*(x_{k+1}) = \pm||\delta_n^*||_\infty = \pm E_n(f), \tag{8.6.1}$$

whereby $\delta_n^*(x) = f(x) - P_n^*(x)$.

The proof of existence and uniqueness of polynomial P_n^* can be found, for example, in [9]. Using the theorem of Weierstrass one can prove that $E_n(f) \to 0$ if $n \to +\infty$.

If for some polynomial of degree n exist $n+2$ points with feature (8.6.1), we say that it has Chebyshev alternation for f. If $\varepsilon_k = (-1)^{k+1}$, (8.6.1) can be expressed as

$$\delta_n^*(x_0) + \varepsilon_k \delta_n^*(x_k) = 0 \quad (k = 1, \ldots, n + 1) \tag{8.6.2}$$

Example 8.6.1. *For function $x \to f(x) = x^{n+1}$ $(|x| \le 1)$ and*

$$\delta_n(x) = x^{n+1} - P_n(x),$$

where $P_n(x) = \sum_{i=0}^{n} a_i x^i$ find the best $\min - \max$ *approximation.*

From the expression for Chebyshev polynomial $T_{n+1}(x) = \cos[(n + 1)arccosx] = \pm 1$, it follows $x_k = -\cos \frac{k\pi}{n+1}$ $(k = 0, 1, \ldots, n + 1)$, where $-1 = x_0 < x_1 < \cdots < x_{n+1} = 1$, we conclude that on $[-1, 1]$ exist $n + 2$ points in which $T_{n+1}(x_k) = (-1)^{n+k+1}$ $(k = 0, 1, \ldots, n + 1)$.

If we put $\delta_n^*(x) = \dfrac{1}{2^n} T_{n+1}(x)$, then $|\delta_n^*(x)| \le \dfrac{1}{2^n}$ $(|x| < 1)$ and $\delta_n^*(x_k) = \dfrac{1}{2^n}(-1)^{n+k+1}$ satisfies (8.6.2).

Based on previous and Theorem (8.6.1) we conclude that the best $\min - \max$ polynomial P_n^* for function $x \to f(x) = x^{n+1}$ $(|x| \le 1)$ can be obtained from equality

$$x^{n+1} - P_n^*(x) = \frac{1}{2^n} T_{n+1}(x).$$

Thus,

$$P_n^*(x) = x^{n+1} - \frac{1}{2^n} T_{n+1}(x).$$

In special case, for $n = 1, 2, 3$ we have

$$x^2 \sim P_1^*(x) = \frac{1}{2}, \quad x^3 \sim P_2^*(x) = \frac{3}{4}x, \quad x^4 \sim P_3^*(x) = x^2 - \frac{1}{8}.$$

Note that approximation of function $x \to x^{n+1}$ $(|x| \le 1)$ in set $\mathcal{P}_n$ using $\min - \max$ approximation, and mean squares approximation with weight function $(1 - x^2)^{-1/2}$, give the same approximating polynomials.

Based on Theorem 8.6.1 the algorithms for determination of best min-max approximation of given function are constructed. One of the most convenient algorithms is Remes algorithm. One version of Remes algorithm can be given in the following way.

1^0 The set of $n+2$ successive points $x_0, x_1, \ldots, x_{n+1}$ on segment $[a,b]$ is to be chosen and coefficients of polynomial P_n and value of E are determined so that

$$f(x_k) - P_n(x_k) = (-1)^k E \quad (k = 0, 1, \ldots, n+1). \tag{8.6.3}$$

2^0 Determine the set of $n+2$ points $\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_{n+1}$, on $[a,b]$ in which $\delta_n(x) = f(x) - P_n(x)$ has successive local extremes with alternative signs, including in this set the point in which the value $|\delta_n(x)|$ has maximal value on $[a,b]$.

3^0 For in advance given accuracy ε the conditions $|\hat{x}_k - x_k| < \varepsilon \quad (k = 0, 1, \ldots, n+1)$ are checked. If at least one of these conditions is not satisfied, one takes $x_k = \hat{x}_k \quad (k = 0, 1, \ldots, n+1)$ and skips to 1^0. When all conditions are satisfied, algorithm ends and polynomial P_n is taken as best min-max approximation P_n^*.

Having in mind that Theorem 8.6.1 (on Chebyshev alternation) can be formulated for some more general types of approximating functions, as, for example, linear approximation function and rational approximating function, Remes algorithm can be applied in this cases too. Note that in case when we have nonlinear approximating function, system of equations (8.6.3) becomes nonlinear and usually has to be solved with Newton-Kantorovich method. In order to shorten the procedure, only a few first steps in iteration are performed.

Very often, for obtaining min-max approximation, the problem is replaced by corresponding discrete mini-max problem. This procedure is given concise in ([1], pp. 121-122). The readers are encouraged to write a code for program realization of given algorithms in arbitrary programming language.

8.7. Packages for approximation of functions

Procedures for developing polynomials for discrete data are very important in engineering practice. The direct fit polynomials, the Lagrange polynomial, and the divided difference polynomial work well for nonequally spaced data. For equally spaced data, polynomials based on differences are recommended.

Procedures for developing least squares approximations for discrete data are also valuable in engineering practice. Least squares approximations are useful for large sets of data and sets of rough data. Least square polynomial approximation is straightforward, for both one independent variable and more than one variable. The least squares normal equations corresponding to polynomial approximating functions are linear, which leads to very efficient solving procedures. For nonlinear approximating functions, the least squares normal equations are nonlinear, which leads to complicated solution procedures. As previously mentioned, convenient mapping of nonlinear approximating function to linear one (i.e. linearization) can solve this problem usually good enough. Least squares polynomial approximation is a straightforward, simple, and accurate procedure for obtaining approximating functions for large sets of data or sets of rough experimental data.

Numerous libraries and software packages are available for approximation of functions, especially for polynomial approximation.

Many commercial software packages contain routines for fitting approximating polynomials. Some of the more prominent packages are `Matlab` and `Mathcad`. More sophisticated packages, such as `IMSL`, `Mathematica`, and `Macsyma` contain also routines for fitting approximating polynomials. The book *Numerical Recipes* ([5]) contains numerous subroutines for fitting approximating polynomials (see Chapter 15, Modelling of Data), and the book *Numerical Methods for Engineers and Scientists.* ([2]) program code for fitting approximating polynomials (see Chapter 4, Polynomial Approximation and Interpolation).

Bibliography (Cited references and further reading)

[1] Milovanović, G.V., *Numerical Analysis II*, Naučna knjiga, Beograd, 1988 (Serbian).

[2] Hoffman, J.D., *Numerical Methods for Engineers and Scientists.* Taylor & Francis, Boca Raton-London-New York-Singapore, 2001.

[3] Milovanović, G.V. and Djordjević, Dj.R., *Programiranje numeričkih metoda na FORTRAN jeziku.* Institut za dokumentaciju zaštite na radu "Edvard Kardelj", Niš, 1981 (Serbian).

[4] Milovanović, G.V., *Numerical Analysis I*, Naučna knjiga, Beograd, 1988 (Serbian).

[4] Stoer, J., and Bulirsch, R., *Introduction to Numerical Analysis*, Springer, New York, 1980.

[5] Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T., *Numerical Recepies - The Art of Scientific Computing.* Cambridge University Press, 1989.

[6] Milovanović, G.V. and Wrigge, S., *Least squares approximation with constraints.* Math. Comp. 46(1986), 551-565.

[7] Milovanović, G.V. and Kovačević, M.A., *Zbirka rešenih zadataka iz numeričke analize.* Naučna knjiga, Beograd, 1985. (Serbian).

[8] Đorđević, Đ.R., Ilić, Z.A., *Realization of Stieltjes procedure and Gauss-Christoffel quadrature formulas in system* `Mathematica`. Zb. rad. Gradj. fak. Niš 17(1996), pp. 29-40.

[9] Todd, J., *Introduction to the Constructive Theory of Functions.* New York, 1963.

[10] Schoenberg,I.J., *Contributions to the problem of approximation of equidistant data by analytic functions.* Quart. Appl. Math. 4(1946), 45-99; 112-141.

[11] Ralston,A., *A First Course in Numerical Analysis.* McGraw-Hill, New York, 1965.

[12] Brass, H., *Approximation durch Teilsummen von Orthogonalpolynomreihen.* In: Numerical Methods of Approximation Theory (L. Collatz, G. Meinarus, H. Werner, eds.), 69-83, ISNM 52, Birkhäuser, Basel, 1980.

[13] Brass, H., *Error estimates for least squares approximation by polynomials.* J. Approx. Theory 41(1984), 345-349.

[14] Collatz, L. and Krabs, W., *Approximationstheorie. Tschebyscheffsche Approximation with Anwendungen.* B.G. Teubner, Stutgart, 1973.

[15] Davis, P.J., *Interpolation and ApproximationBrass.* Blaisdell Publ. Comp., New York, 1963.

[16] Feinerman, R.P. & Newman, D.J., *Polynomial Approximation.* Williams & Wilkins, Baltimore, 1974.

[17] Fox, L. & Parker, I.B., *Chebyshev Polynomials in Numerical Analysis.* Oxford Univ. Press, 1972.

[18] Hildebrand, F.B., *Introduction to Numerical Analysis.* Mc.Graw-Hill, New York, 1974.

[19] Rivlin, T.J., *An Introduction to the Approximation od Functions.* Dover Publications, Inc., New York, 1981.

[20] Stoer, J., and Bulirsch, R., *Introduction to Numerical Analysis*, Springer, New York, 1980.

[21] *IMSL Math/Library Users Manual* , IMSL Inc., 2500 City West Boulevard, Houston TX 77042

[22] *NAG Fortran Library*, Numerical Algorithms Group, 256 Banbury Road, Oxford OX27DE, U.K., Chapter F02.