Выпишем формулу (2.12.5) при $q = \frac{1}{2}$. Получим $P(x) = \frac{y_0 + y_1}{2} - \frac{\Delta^2 y_{-1} + \Delta^2 y_0}{16} +$ $+ \frac{3(\Delta^4 y_{-2} + \Delta^4 y_{-1})}{256}$. По имеющейся таблице и по только что написанной формуле можно получить значение функции только для $x = 2.25$. Тогда $q = \frac{2.25 - 2.20}{0.1} = 0.5$ и $f(2.25) = 2.0060 - 0.0020 + 0.0000 = 2.0040$. По данной таблице можно получить еще два значения функции для $x = 2.15$ и $x = 2.35$, так как вклад последнего члена формулы для $P(x)$ меньше $4 \cdot 10^{-6}$, поэтому для $x = 2.15$ и $x = 2.35$ четвертые разности не понадобятся (их нет для значений $x = 2.15$ и $x = 2.35$). Структура формулы для $P(x)$ для этих значений не изменится, только разности, входящие в формулу, сдвинутся по таблице для $x = 2.15$ вверх, а для $x = 2.35$ - вниз на одну позицию (для $x = 2.15$ они обведены пунктиром):

$$f(2.15) = \frac{1.9141 + 1.7455}{2} - \frac{0.0137 + 0.0151}{16} = 1.8298 - 0.0018 = 1.8280 \text{ и}$$

$$f(2.35) = \frac{2.2981 + 2.0978}{2} - \frac{0.0166 + 0.0183}{16} = 2.1980 - 0.0022 = 2.1958.$$

3. МЕТОД НАИМЕНЬШИХ КВАДРАТОВ И СПЕЦИАЛЬНЫЕ ИНТЕРПОЛЯЦИОННЫЕ МНОГОЧЛЕНЫ

3.1. Постановка задачи и вывод формул метода наименьших квадратов

Задача наименьших квадратов возникает в самых различных областях науки и техники, например, к ней приходят при статистической обработке экспериментальных данных. Пусть функция $y = f(x)$ задана таблицей приближенных значений $y_i \approx f(x_i)$, $i = \overline{0, n}$, полученных с ошибками $\varepsilon_i = y_i^\circ - y_i$, где $y_i^\circ = f(x_i)$. Предположим, что для аппроксимации функции $y = f(x)$ используется линейная модель: $y = \Phi_m(x) = a_0\varphi_0(x) +$ $+ a_1\varphi_1(x) + \ldots + a_m\varphi_m(x)$, где $\varphi_0(x), \varphi_1(x), \ldots, \varphi_m(x)$ - заданные базисные функции, $a_0, a_1, \ldots, a_m$ - параметры модели, являющиеся одновременно коэффициентами обобщенного многочлена. Часто используется одна из наиболее простых моделей $y = P_m(x) = a_0 + a_1 x + \ldots + a_m x^m$ - полиномиальная модель.

В случае, когда уровень неопределенности исходных данных высок, нет смысла требовать точного совпадения значений обобщенного многочлена $\Phi_m(x)$ в точках x_i с заданными значениями y_i, то есть использовать интерполяцию. Кроме того, при интерполяции происходит повторение ошибок наблюдений, в то время как при обработке экспериментальных данных желательно сглаживание ошибок. Тем не менее нужно потребовать, чтобы

$$\begin{cases} a_0\varphi_0(x_0) + a_1\varphi_1(x_0) + \ldots + a_m\varphi_m(x_0) \approx y_0, \\ a_0\varphi_0(x_1) + a_1\varphi_1(x_1) + \ldots + a_m\varphi_m(x_1) \approx y_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0\varphi_0(x_n) + a_1\varphi_1(x_n) + \ldots + a_m\varphi_m(x_n) \approx y_n. \end{cases} \tag{3.1.1}$$

Эта же система в матричной форме имеет вид $P\overline{a} \approx \overline{y}$. (3.1.2)

Существуют разные дополнительные критерии, позволяющие решить эту систему, так как в общем случае при $m < n$ она, вообще говоря, несовместна. Выбор $\overline{a}$, позволяющий наилучшим образом удовлетворить (3.1.2) в методе наименьших квадратов, состоит в том минимизируется среднее квадратическое уклонение

$$\delta(\Phi_m, y) = \sqrt{\frac{1}{n+1}\sum_{i=0}^{n}(\Phi_m(x_i) - y_i)^2}. \tag{3.1.3}$$

Итак, линейная задача метода наименьших квадратов состоит в следующем. Надо найти обобщенный многочлен $\Phi_m(x)$, для которого среднеквадратическое уклонение $\delta(\Phi_m, y) \Rightarrow \min$. Этот многочлен называется *многочленом наилучшего среднего квадратического приближения*. Так как набор функций $\{\varphi_i(x)\}_0^m$ всегда заранее определен, задача заключается в нахождении вектора $\bar{a} = (a_0^*, a_1^*, ... a_m^*)^T$ при условии $\delta(\Phi_m, y) \Rightarrow \min$. Для решения нашей задачи воспользуемся общим приемом дифференциального исчисления, а именно выпишем необходимые условия экстремума функции нескольких переменных (приравняем частные производные нулю):

$$\begin{cases} \dfrac{\partial S}{\partial a_k} = 0, \ k = 0,1,...,m, \text{ где} \\ S = \sum\limits_{i=0}^{n} \sum\limits_{j=0}^{m} \left(a_j \varphi_j(x_i) - y_i\right)^2. \end{cases} \tag{3.1.4}$$

Тогда получим $\dfrac{\partial S}{\partial a_k} = 2\sum\limits_{i=0}^{n} \sum\limits_{j=0}^{m} \left(a_j \varphi_j(x_i) - y_i\right) \cdot \varphi_k(x_i) - 0, \ k = \overline{0, m}.$ Изменим в первом слагаемом порядок суммирования:

$$\sum_{j=0}^{m} \left(\sum_{i=0}^{n} \varphi_j(x_i) \cdot \varphi_k(x_i) \right) a_j = \sum_{i=0}^{n} y_i \varphi_k(x_i), \ k = \overline{0, m}. \tag{3.1.5}$$

Уравнение (3.1.5) называется нормальной системой метода наименьших квадратов.

Если вернуться к обозначениям формулы (3.1.2), то, как нетрудно видеть, систему (3.1.5) можно записать в виде

$$P_{m\times n}^T P_{n\times m} = P_{m\times m}^T \bar{y}. \tag{3.1.6}$$

Матрица $P_{m\times n}^T P_{n\times m} = \Gamma_{m\times m}$ называется матрицей Грама*. Если еще ввести вектор $\bar{b} = P^T \bar{y}$, то система (3.1.6) перепишется в виде $\Gamma \bar{a} = \bar{b}$ - система линейных уравнений относительно вектора $\bar{a}$. Можно показать, что если среди точек $x_0, x_1, ..., x_n$ нет совпадающих и $m \le n$, то определитель системы (3.1.6) отличен от нуля, и, следовательно, эта система имеет единственное решение: $a_0 = a_0^*, a_1 = a_1^*, ..., a_m = a_m^*$. Обобщенный полином с такими коэффициентами будет обладать минимальным средним квадратическим отклонением $\delta(\Phi_m, y)$.

Если $m = n$, то обобщенный многочлен, если система функций $\{\varphi_i\}_{i=\overline{0,m}}$ степенная, совпадает с полиномом Лагранжа для системы точек $x_0, x_1, ..., x_n$, причем $S_{\min} = 0$. При $m < n$ построение такого точного интерполяционного многочлена невозможно. Таким образом, аппроксимация функций представляет собой более общий процесс, чем интерполирование.

Если $\{\varphi_i(x)\}_0^m = \{x^i\}_0^m$, то нормальная система (3.1.5) принимает следующий вид:

$$\begin{cases} S = \sum\limits_{i=0}^{n} \sum\limits_{j=0}^{m} \left(a_j x_i^j - y_i\right)^2, \\ \dfrac{\partial S}{\partial a_k} = 2\sum\limits_{i=0}^{n} \sum\limits_{j=0}^{m} \left(a_j x_i^j - y_i\right) x_i^k = 0, \\ \sum\limits_{j=0}^{m} \left(\sum\limits_{i=0}^{n} x_i^{j+k} \right) a_j = \sum\limits_{i=0}^{n} y_i x_i^k, \ k = \overline{0, m}. \end{cases} \tag{3.1.7}$$

Запишем систему (3.1.7) в развернутом виде в двух наиболее простых случаях при $m = 1$ и $m = 2$. В случае, когда приближение осуществляется многочленом первой степени $P_1(x) = a_0 + a_1 x$, уравнения метода наименьших квадратов имеют следующий вид:

* Иорген Педерсен Грам (1850-1916) - датский математик.

$$S=\sum_{i=0}^{n}\sum_{j=0}^{1}\left(a_j x_i^j - y_i\right)^2=\sum_{i=0}^{n}\left(a_0+a_1x_i-y_i\right)^2 \Rightarrow \min,$$

$$\frac{\partial S}{\partial a_k}=2\sum_{i=0}^{n}\left(a_0+a_1x_i-y_i\right)\cdot x_i^k=0,\ k=0{,}1 \text{ или}$$

$$\begin{cases}\dfrac{\partial S}{\partial a_0}=2\sum\limits_{i=0}^{n}\left(a_0+a_1x_i-y_i\right)\cdot 1=0,\\ \dfrac{\partial S}{\partial a_1}=2\sum\limits_{i=0}^{n}\left(a_0+a_1x_i-y_i\right)\cdot x_i=0.\end{cases}$$

$$\begin{cases}a_0(n+1)+a_1\sum\limits_{i=0}^{n}x_i=\sum\limits_{i=0}^{n}y_i,\\ a_0\sum\limits_{i=0}^{n}x_i+a_1\sum\limits_{i=0}^{n}x_i^2=\sum\limits_{i=0}^{n}x_iy_i\end{cases} \tag{3.1.8}$$

- нормальная система для $m=1$ в развернутом виде. Пусть теперь $m=2$. Аналогично получим $P_2(x)=a_0+a_1x+a_2x^2$.

$$S=\sum_{i=0}^{n}\sum_{j=0}^{2}\left(a_j x_i^j - y_j\right)^2=\sum_{i=0}^{n}\left(a_0+a_1x+a_2x^2-y_i\right)^2 \Rightarrow \min.$$

$$\begin{cases}\dfrac{\partial S}{\partial a_0}=2\sum\limits_{i=0}^{n}\left(a_0+a_1x+a_2x^2-y_i\right)\cdot 1=0,\\ \dfrac{\partial S}{\partial a_1}=2\sum\limits_{i=0}^{n}\left(a_0+a_1x+a_2x^2-y_i\right)\cdot x_i=0,\\ \dfrac{\partial S}{\partial a_2}=2\sum\limits_{i=0}^{n}\left(a_0+a_1x+a_2x^2-y_i\right)\cdot x_i^2=0.\end{cases}$$

$$\begin{cases}a_0(n+1)+a_1\sum\limits_{i=0}^{n}x_i+a_2\sum\limits_{i=0}^{n}x_i^2=\sum\limits_{i=0}^{n}y_i,\\ a_0\sum\limits_{i=0}^{n}x_i+a_1\sum\limits_{i=0}^{n}x_i^2+a_2\sum\limits_{i=0}^{n}x_i^3=\sum\limits_{i=0}^{n}x_iy_i,\\ a_0\sum\limits_{i=0}^{n}x_i^2+a_1\sum\limits_{i=0}^{n}x_i^3+a_2\sum\limits_{i=0}^{n}x_i^4=\sum\limits_{i=0}^{n}x_i^2y_i\end{cases} \tag{3.1.9}$$

- нормальная система для $m=2$ в развернутом виде для квадратичного сглаживания.

Метод вычисления параметров $a_0^*, a_1^*, \ldots, a_m^*$ с помощью решения нормальной системы кажется весьма привлекательным. Действительно, задача сводится к стандартной системе линейных алгебраический уравнений с квадратной матрицей. Однако вычислительная практика показывает, что без специального выбора базисных функций $\{\varphi_i(x)\}_0^m$ уже при $m\geq 5$ нормальная система обычно оказывается плохо обусловленной. Причина в том, что система базисных функций, будучи формально независимой, на практике часто близка к линейно зависимой. Особенно этим «грешит» система степенных функций $1, x, x^2, \ldots, x^m$, широко применяемая при аппроксимации алгебраическими многочленами. Лучший результат получается, если использовать систему ортогональных на отрезке $[a, b]$ функций. Пример такой системы на $[-1, 1]$ дает система многочленов Чебышева $T_0(x), T_1(x), \ldots, T_m(x)$.

В настоящее время в вычислительной практике нормальная система, как правило, не используется. Применяются другие, более надежные методы, например метод сингулярного разложения матрицы P.

Пример. Пусть функция $y = f(x)$ задана следующей таблицей:

x	0.78	1.56	2.34	3.12	3.81
y	2.50	1.20	1.12	2.25	4.28

Используя метод наименьших квадратов, аппроксимируем ее многочленами первой и второй степени и найдем соответствующие средние квадратические уклонения δ_1 и δ_2.

Вычисления, которые нужно провести, расположим по схеме, приведенной в такой таблице:

x^0	x^1	x^2	x^3	x^4	y	xy	x^2y
1	0.78	0.608	0.475	0.370	2.50	1.950	1.521
1	1.56	2.434	3.796	5.922	1.20	1.872	2.920
1	2.34	5.476	12.813	29.982	1.12	2.621	6.133
1	3.12	9.734	30.371	94.759	2.25	7.020	21.902
1	3.81	14.516	55.306	210.717	4.28	16.307	62.129
$\sum$ 5	11.61	32.768	102.761	341.750	11.35	29.770	94.605

а) Линейная модель

$$\begin{cases} a_0(n+1) + a_1\sum_{i=0}^{n} x_i = \sum_{i=0}^{n} y_i, \\ a_0\sum_{i=0}^{n} x_i + a_1\sum_{i=0}^{n} x_i^2 = \sum_{i=0}^{n} x_i y_i. \end{cases}$$

$$\begin{cases} 5a_0 + 11.61a_1 = 11.35, \\ 11.61a_0 + 32.768a_1 = 29.770, \end{cases} \Rightarrow \begin{cases} a_0 + 2.322a_1 = 2.27, \\ 5.810a_1 = 3.415, \\ a_1 = 0.588, \\ a_0 = 0.905. \end{cases}$$

Таким образом, линейная модель имеет вид $y = 0.905 + 0.588x$.

б) Квадратичная модель

$$\begin{cases} a_0(n+1) + a_1\sum_{i=0}^{n} x_i + a_2\sum_{i=0}^{n} x_i^2 = \sum_{i=0}^{n} y_i, \\ a_0\sum_{i=0}^{n} x_i + a_1\sum_{i=0}^{n} x_i^2 + a_2\sum_{i=0}^{n} x_i^3 = \sum_{i=0}^{n} x_i y_i, \\ a_0\sum_{i=0}^{n} x_i^2 + a_1\sum_{i=0}^{n} x_i^3 + a_2\sum_{i=0}^{n} x_i^4 = \sum_{i=0}^{n} x_i^2 y_i. \end{cases}$$

$$\begin{cases} 5a_0 + 11.61a_1 + 32.768a_2 = 11.35, \\ 11.61a_0 + 32.768a_1 + 102.761a_2 = 29.770, \\ 32.768a_0 + 102.761a_1 + 341.750a_2 = 94.605. \end{cases} \Rightarrow \begin{cases} a_0 + 2.322a_1 + 6.564a_2 = 2.27, \\ 5.810a_1 + 26.553a_2 = 3.415, \\ 26.674a_1 + 126.661a_2 = 20.222. \end{cases}$$

$$\begin{cases} a_0 + 2.322a_1 + 6.564a_2 = 2.27, \\ a_1 + 4.570a_2 = 0.588, \\ 4.761a_2 = 4.538. \end{cases} \Rightarrow \begin{cases} a_2 = 0.953, \\ a_1 = -3.767, \\ a_0 = 4.762. \end{cases}$$

Отсюда $y = 4.762 - 3.767x + 0.953x^2$ - вид квадратичной модели. Обе модели значительно отличаются друг от друга. Сравним исходные данные для $y = f(x)$ с соответствующими значениями y^*, полученными из обеих моделей, и вычислим δ_1 и δ_2.

x	y	y_1^*	$y_1^* - y$	$(y_1^* - y)^2$	y_2^*	$y_2^* - y$	$(y_2^* - y)^2$
0.78	2.50	1.364	-1.136	1.290	2.404	-0.096	0.009
1.56	1.20	1.822	0.622	0.387	1.204	0.004	0.000
2.34	1.12	2.281	1.161	1.350	1.165	0.045	0.002
3.12	2.25	2.740	0.490	0.240	2.286	0.036	0.001
3.81	4.28	3.145	1.135	1.290	4.244	-0.036	0.001
			Σ	4.557		Σ	0.013

Таким образом, $\delta_1 = \sqrt{\frac{1}{5} \cdot 4.557} = 0.955$, $\delta_2 = \sqrt{\frac{1}{5} \cdot 0.013} = 0.051$. Следовательно, данным для $y = f(x)$ в исходной таблице очень хорошо соответствует квадратичная модель. Линейная модель не адекватна исходным данным и должна быть отвергнута.

3.2. Лабораторная работа № 4. Аппроксимация функции по методу наименьших квадратов

Очень часто при анализе эмпирических данных необходимо найти явную функциональную зависимость между двумя величинами x и y, полученными в результате измерений. Поскольку опытные данные всегда содержат ошибки, то строить интерполяционный многочлен $y = P_n(x)$ не рационально, так как при интерполяции ошибки повторяются. Желательно по возможности сгладить и минимизировать ошибки наблюдений. Этот результат достигается построением многочлена наилучшего среднего квадратического приближения по методу наименьших квадратов.

Итак, если $y = f(x)$ аппроксимируется многочленом вида $P_m(x) = \sum_{j=0}^{m} a_j x^j$, так что система базисных функций имеет вид $\{\varphi_i(x)\}_0^m = \{x^i\}_0^m$, то неизвестные коэффициенты многочлена $P_m(x)$ по методу наименьших квадратов определяются из решения системы (3.1.7).

В подразделе 3.1 описан пример «ручного» вычисления коэффициентов линейной и квадратичной модели по методу наименьших квадратов. Решим аналогичную задачу средствами пакета Mathcad различными способами. Сформируем вначале вектора исходных данных.

В алгебре матриц в среде Mathcad доступны несколько очень удобных встроенных функций, например, submatrix, stack и augment. Функция $\mathrm{submatrix}(A, m, n, k, l)$ извлекает из матрицы A подматрицу, содержащуюся в A со строки m по строку n и со столбца с номером k по номер l. Функции stack и augment, наоборот, формируют одну матрицу из двух. После работы $\text{stack}(A, B)$ получается массив, сформированный расположением A над B, при этом матрицы A и B должны иметь одинаковое число столбцов. Функция $\text{augment}(A, B)$ располагает матрицы A и B рядом, B справа от A; эти матрицы должны иметь одинаковое число строк.

Введем с клавиатуры

$$\text{ORIGIN} := 1 \quad n := 20 \quad i := 1..n \quad x_i := \frac{i}{10}$$

$$y1 := \begin{pmatrix} -3.97 & -5.83 \\ -4.07 & -6.06 \\ -4.04 & -6.40 \\ -4.30 & -6.83 \\ -4.27 & -7.54 \\ -4.54 & -7.68 \\ -4.79 & -8.36 \\ -5.07 & -8.91 \\ -5.30 & -9.34 \\ -5.51 & -9.98 \end{pmatrix} \quad \begin{array}{l} y2 := \text{submatrix}(y1,1,10,2,2) \\ y1 := \text{submatrix}(y1,1,10,1,1) \end{array} \quad y1 = \begin{pmatrix} -3.97 \\ -4.07 \\ -4.04 \\ -4.30 \\ -4.27 \\ -4.54 \\ -4.79 \\ -5.07 \\ -5.30 \\ -5.51 \end{pmatrix} \quad y2 = \begin{pmatrix} -5.83 \\ -6.06 \\ -6.40 \\ -6.83 \\ -7.54 \\ -7.68 \\ -8.36 \\ -8.91 \\ -9.34 \\ -9.98 \end{pmatrix}$$

$$y := \text{stack}(y1, y2) \quad y3 := \text{augment}(y1, y2) \quad y4 := \text{augment}(y2, y1)$$

$$y = \begin{pmatrix} -3.97 \\ -4.07 \\ -4.04 \\ -4.30 \\ -4.27 \\ -4.54 \\ -4.79 \\ -5.07 \\ -5.30 \\ -5.51 \\ -5.83 \\ -6.06 \\ -6.40 \\ -6.83 \\ -7.54 \\ -7.68 \\ -8.36 \\ -8.91 \\ -9.34 \\ -9.98 \end{pmatrix} \quad y3 = \begin{pmatrix} -3.97 & -5.83 \\ -4.07 & -6.06 \\ -4.04 & -6.40 \\ -4.30 & -6.83 \\ -4.37 & -7.54 \\ -4.54 & -7.68 \\ -4.79 & -8.36 \\ -5.07 & -8.91 \\ -5.30 & -9.34 \\ -5.51 & -9.98 \end{pmatrix} \quad y4 = \begin{pmatrix} -5.83 & -3.97 \\ -6.06 & -4.07 \\ -6.40 & -4.04 \\ -6.83 & -4.30 \\ -7.54 & -4.27 \\ -7.68 & -4.54 \\ -8.36 & -4.79 \\ -8.91 & -5.07 \\ -9.34 & -5.30 \\ -9.98 & -5.51 \end{pmatrix}$$

$$m := \text{cols}(y3) \quad m = 2 \quad n1 := \text{rows}(y4) \quad n1 = 10$$

$$\min(y) = -9.98 \quad \max(y) = -3.97 \quad \text{length}(y1) = 10 \quad \text{last}(y) = 20$$

Функции $\text{cols}(A)$ и $\text{rows}(A)$ возвращают число столбцов и строк матрицы A, $\min(A)$ и $\max(A)$ соответственно наименьшее и наибольшее значение элементов в A, $\text{length}(a)$ - число элементов в векторе $\bar{a}$, $\text{last}(a)$ - индекс последнего элемента в векторе $\bar{a}$ с учетом значения переменной ORIGIN.

Построим линейную и квадратичную модель по формулам (3.1.8) и (3.1.9). Для этого вычислим следующие величины. Введем еще одну предопределенную переменную пакета Mathcad $TOL := 10^{-6}$. Она определяет допустимую погрешность для различных алгоритмов аппроксимации, интегрирования, решения уравнений и так далее. По умолчанию $TOL = 10^{-3}$. Вычислим следующие величины:

$$x1 := \sum_{i=1}^{n} x_i \quad x1 = 21 \quad x2 := \sum_{i=1}^{n} (x_i)^2 \quad x2 = 28.7 \quad x3 := \sum_{i=1}^{n} (x_i)^3 \quad x3 = 44.1$$

$$x4 := \sum_{i=1}^{n} (x_i)^4 \quad x4 = 72.2666 \quad xy := \sum_{i=1}^{n} x_i \cdot y_i \quad xy = -149.812$$

$$x2y := \sum_{i=1}^{n} (x_i)^2 \cdot y_i \quad x2y = -222.6264 \quad y11 := \sum_{i=1}^{n} y_i \quad y11 = -122.79$$

$$A1 := \begin{pmatrix} 20 & 21 \\ 21 & 28.7 \end{pmatrix} \quad A1^{-1} = \begin{pmatrix} 0.215789 & -0.157895 \\ -0.157895 & 0.150376 \end{pmatrix} \quad b1 := \begin{pmatrix} -122.79 \\ -149.812 \end{pmatrix}$$

$$a1 := A1^{-1} \cdot b1 \quad a1 = \begin{pmatrix} -2.842263 \\ -3.140226 \end{pmatrix}$$

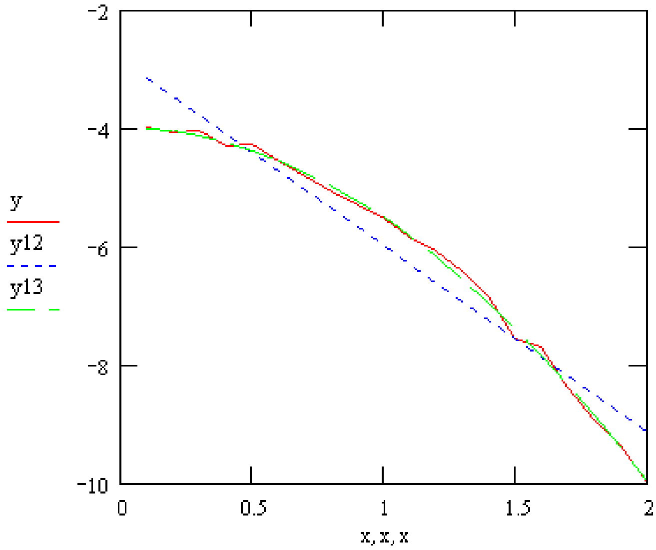
$$A := \begin{pmatrix} 20 & 21 & 28.7 \\ 21 & 28.7 & 44.1 \\ 28.7 & 44.1 & 72.2666 \end{pmatrix} \quad A^{-1} = \begin{pmatrix} 0.443509 & -1.078947 & 0.438596 \\ -1.078947 & 2.662338 & -1.196172 \\ 0.438596 & -1.196172 & 0.569606 \end{pmatrix}$$

$$b := \begin{pmatrix} -122.79 \\ -149.812 \\ -222.6264 \end{pmatrix} \quad a2 := A^{-1} \cdot b \quad a2 = \begin{pmatrix} -3.969237 \\ -0.066661 \\ -1.463602 \end{pmatrix}$$

$$i := 1..n \quad y12_i := a1_1 + a1_2 \cdot x_i \quad y13_i := a2_1 + a2_2 \cdot x_i + a2_3 \cdot (x_i)^2$$

$$r1_i := (y_i - y12_i)^2 \quad r2_i := (y_i - y13_i)^2$$

$$\delta 1 := \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} r1_i} \quad \delta 2 := \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} r2_i} \quad \delta 1 = 0.442 \quad \delta 2 = 0.085$$



Коэффициенты системы нормальных уравнений линейной модели, то есть системы (3.1.8), находятся в матрице $A1$, коэффициенты квадратичной модели (3.1.9) - в матрице A. Решение обеих систем линейных уравнений произведено с помощью обратной матрицы. Вектор $\overline{a1}$ содержит коэффициенты линейной модели, вектор $\overline{a2}$ - квадратичной.

Далее вычисляются невязки по обеим моделям и находятся средние квадратические ошибки $\delta 1$ и $\delta 2$. Видно, что исходным данным хорошо удовлетворяет квадратичная модель $y_i = -3.969237 - 0.066661x_i - 1.463602x_i^2$. Этот факт отчетливо виден и на приведенном графике.

Mathcad не назывался бы математическим пакетом, если бы не умел решать алгебраические системы различными, в том числе и более эффективными способами. Одним из таких способов является конструкция **Given – Find**. Это две команды: **Given** (Дано) и **Find** (Найти). Сначала задается какое-нибудь начальное приближение, например, для квадратичной модели

$$\begin{pmatrix} x5 \\ y5 \\ z \end{pmatrix} := \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

затем за ключевым словом **Given** нужно записать анализируемую систему, связывая левые и правые части уравнений знаком «эквивалентно» (жирным знаком «равно» из панели равенств и отношений или же нажимая сразу обе клавиши $\lfloor \mathrm{Ctrl} \rfloor \lfloor = \rfloor$), после этого должно идти второе ключевое слово **Find**. Эта функция возвращает решение анализируемой системы:

$$\text{Given} \quad 20 \cdot x5 + 21 \cdot y5 + 28.7 \cdot z = -122.79$$

$$21 \cdot x5 + 28.7 \cdot y5 + 44.1 \cdot z = -149.812$$

$$28.7 \cdot x5 + 44.1 \cdot y5 + 72.26666 \cdot z = -222.6264$$

$$\text{Find}(x5, y5, z) = \begin{pmatrix} -3.969237 \\ -0.066661 \\ -1.463602 \end{pmatrix}$$

Неудобство применения пары **Given – Find** в том, что решаемая система уравнений должна быть записана в скалярной форме. Вместо функции **Find** можно использовать пару **Given – MinErr**. Функция $\text{MinErr}(x, y, \ldots)$ дает решение системы уравнений, которое приводит к минимальным невязкам. Число неизвестных системы должно быть равно числу аргументов функции **MinErr**. В нашем случае

$$\begin{pmatrix} xx \\ yy \\ zz \end{pmatrix} := \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{Given} \quad 20 \cdot xx + 21 \cdot yy + 28.7 \cdot zz = -122.79$$

$$21 \cdot xx + 28.7 \cdot yy + 44.1 \cdot zz = -149.812$$

$$28.7 \cdot xx + 44.1 \cdot yy + 72.26666 \cdot zz = -222.6264$$

$$\text{MinErr}(xx, yy, zz) = \begin{pmatrix} -3.969237 \\ -0.066661 \\ -1.463602 \end{pmatrix}.$$

Наконец, для решения линейных систем алгебраических уравнений можно использовать встроенную функцию **lsolve**. Она возвращает вектор решения системы, записанной в матричном виде:

$$a22 := \text{lsolve}(A, b) \quad a22 = \begin{pmatrix} -3.969237 \\ -0.066661 \\ -1.463602 \end{pmatrix}$$

Заметим, что функцию **lsolve** можно использовать в программируемых конструкциях, тогда как пары **Given – Find** и **Given – MinErr** этого не допускают.

Приведем в заключение подпрограмму, реализующую вычисления по формуле (3.1.7) в общем случае коэффициентов сглаживающего многочлена заданной степени. Все операторы этой подпрограммы легко отождествляются с той или иной частью формулы (3.1.7). Параметры подпрограммы: x, y - вектора исходных данных, n - число точек сетки таблично заданной функции, m - требуемая степень сглаживающего многочлена. В результате работы подпрограмма МНК выдает вектор коэффициентов многочлена $P_m(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_m x^m$, записанных в следующем порядке: $a_0, a_1, a_2, \ldots, a_{m-1}, a_m, \delta$. Последняя $m+2$ компонента вектора результата содержит среднюю квадратическую ошибку представления исходных табличных данных построенным сглаживающим многочленом:

```
MHK(x,y,n,m) := | for k ∈ 1..m+1
                |   | for j ∈ 1..m+1
                |   |
                |   |   A_{j,k} ← Σ_{l=1}^{n} (x_l)^(j+k-2)
                |   |
                |   | b_k ← Σ_{l=1}^{n} y_l·(x_l)^(k-1)
                | a ← lsolve(A,b)
                | for i ∈ 1..n
                |   | y1_i ← 0
                |   | for j ∈ 1..m+1
                |   |   y1_i ← y1_i + a_j·(x_i)^(j-1)
                | r ← Σ_{i=1}^{n} (y_i - y1_i)^2
                | a_{m+2} ← sqrt(r/n)
                | a
```

Для нашего примера

$$a23 := MHK(x, y, 20, 1) \quad a23 = \begin{pmatrix} -2.842263 \\ -3.140226 \\ 0.441879 \end{pmatrix}$$

$$a24 := MHK(x, y, 20, 2) \quad a24 = \begin{pmatrix} -3.969237 \\ -0.066661 \\ -1.463603 \\ 0.084973 \end{pmatrix}$$

$$a25 := MHK(x, y, 20, 3) \quad a25 = \begin{pmatrix} -3.904312 \\ -0.397947 \\ -1.078670 \\ -0.122201 \\ 0.083012 \end{pmatrix}$$

$$a26 := MHK(x, y, 20, 4) \quad a26 = \begin{pmatrix} -3.909030 \\ -0.361037 \\ -1.153360 \\ -0.067807 \\ -0.012951 \\ 0.083006 \end{pmatrix}$$

Видно, что для исходной таблично заданной функции многочленом наилучшего приближения является уже полученный ранее многочлен второй степени $P_2(x) = -3.969237 - 0.066661x - 1.463602x^2$. Дальнейшее усложнение модели (повышение степени многочлена) практически не изменяет среднюю квадратическую ошибку и, следовательно, не является оправданным.

Задание № 1. По методу наименьших квадратов аппроксимировать таблично заданную функцию $y = f(x)$ многочленом наилучшего среднеквадратического приближения $(1 \le m \le 4)$.

Т а б л и ц а 1

x_i	$y_i = f(x_i)$									
	Номера вариантов									
	1	2	3	4	5	6	7	8	9	10
0.1	3.15	20.95	2.20	6.00	3.13	0.07	12.28	-9.10	9.84	3.09
0.2	3.04	20.51	2.18	7.04	3.19	0.17	12.53	-9.21	10.01	3.31
0.3	3.02	21.96	1.87	7.21	3.17	0.21	12.50	-8.99	11.10	3.72
0.4	2.97	21.83	1.85	7.40	3.52	0.31	12.53	-8.95	12.16	3.77
0.5	2.87	21.79	1.77	7.20	3.62	1.10	12.75	-9.13	13.05	3.78
0.6	2.98	22.72	1.62	7.70	3.72	1.09	12.85	-9.23	14.35	3.97
0.7	2.81	25.80	1.57	7.36	4.03	1.12	12.77	-9.21	15.19	4.00
0.8	2.70	27.33	1.27	7.61	4.39	-0.37	12.76	-9.43	15.50	4.51
0.9	2.66	28.21	1.05	7.56	4.72	-0.22	12.73	-9.57	15.74	4.43
1.0	2.50	30.45	0.68	7.50	4.85	-0.48	12.85	-9.44	16.03	4.58
1.1	2.60	30.37	0.55	7.51	5.12	-0.84	12.51	-9.44	16.56	4.58
1.2	2.36	34.51	-0.10	7.53	5.38	-0.93	12.34	-9.83	17.49	4.54
1.3	2.09	36.29	-0.41	7.45	5.96	-1.15	12.22	-9.78	17.79	4.82
1.4	2.07	38.53	-1.00	7.27	6.40	-1.44	11.84	-9.81	18.03	4.90
1.5	2.01	41.90	-1.19	7.20	6.58	-1.90	11.67	-10.06	18.82	4.77
1.6	1.81	44.52	-1.56	7.25	7.09	-2.25	11.27	-10.41	19.50	4.81
1.7	1.53	48.91	-2.08	7.35	7.32	-2.65	11.06	-10.40	20.28	5.00
1.8	1.64	50.68	-2.61	6.97	7.94	-3.06	10.73	-10.70	21.21	4.97
1.9	1.29	56.36	-3.37	7.20	8.47	-3.66	10.35	-10.96	22.63	5.08
2.0	1.11	59.14	-3.86	7.06	9.00	-4.01	10.09	-11.91	22.90	5.08

Т а б л и ц а 2

x_i	$y_i = f(x_i)$									
	Номера вариантов									
	11	12	13	14	15	16	17	18	19	20
0.1	8.15	-6.90	0.17	3.30	1.04	0.08	3.09	-0.86	0.00	-0.65
0.3	8.41	-7.01	0.07	2.49	1.47	0.14	3.25	-0.77	-0.09	-1.00
0.5	8.58	-7.19	0.17	3.02	1.78	0.37	3.48	-0.56	-0.26	-0.87
0.7	8.84	-7.11	0.05	3.27	2.01	0.36	3.47	-0.46	-0.20	-0.89
0.9	9.28	-7.31	0.12	3.43	2.19	0.44	3.55	-0.28	-0.29	-0.75
1.1	9.46	7.78	0.00	3.70	2.60	0.48	3.59	-0.24	-0.14	-0.59
1.3	10.02	-7.64	0.01	3.70	2.93	0.27	3.28	-0.36	-0.26	-0.44
1.5	10.11	-7.85	-0.05	3.85	3.22	0.39	3.50	-0.43	-0.45	-0.61
1.7	10.61	-8.18	-0.21	3.89	3.50	0.50	3.61	-0.56	-0.43	-0.17
1.9	11.03	-8.39	-0.50	3.98	4.01	0.48	3.59	-0.59	-0.71	0.13
2.1	11.34	-8.79	-0.50	4.02	4.22	0.69	3.80	-0.70	-0.70	0.53
2.3	11.86	-9.02	-0.86	4.21	4.71	0.50	3.61	-1.01	-1.00	0.67
2.5	12.33	-9.48	-1.24	4.22	5.23	0.31	3.42	-1.03	-1.01	1.00
2.7	12.81	-9.93	-1.47	4.37	5.78	0.37	3.48	-1.47	-1.17	1.34
2.9	13.21	-10.26	-1.79	4.36	6.27	0.43	3.54	-1.68	-1.39	1.49
3.1	13.67	-10.91	-2.25	4.39	6.75	0.33	3.44	-1.93	-1.22	1.81

3.3	14.23	-11.41	-2.55	4.54	7.16	0.31	3.42	-2.28	-1.43	2.37
3.5	14.68	-11.91	-3.18	4.33	7.76	0.09	3.20	-2.53	-1.81	2.72
3.7	15.35	-12.30	-3.60	4.54	8.30	0.08	3.19	-2.93	-1.84	3.03
3.9	15.93	-13.00	-3.93	4.53	9.00	0.03	3.14	-3.07	-1.99	3.51

Т а б л и ц а 3

x_i	$y_i = f(x_i)$									
	Номера вариантов									
	21	22	23	24	25	26	27	28	29	30
0.15	0.16	1.89	-1.92	1.10	-2.80	0.00	4.01	0.12	4.13	2.97
0.30	0.02	2.07	-1.60	1.20	-2.66	0.01	4.06	0.31	4.11	3.07
0.45	0.28	2.30	-1.57	1.18	-2.36	0.24	3.88	0.48	3.87	3.04
0.60	0.42	2.26	-1.41	1.14	-2.41	0.74	3.98	0.45	3.74	3.30
0.75	0.31	2.34	-1.36	1.17	-2.13	1.02	4.36	0.84	3.85	3.27
0.90	0.41	2.66	-0.97	1.00	-1.82	1.31	4.18	0.73	3.71	3.54
1.05	0.42	2.88	-0.59	0.99	-1.74	1.53	4.16	0.77	3.53	3.79
1.20	0.36	2.85	-0.71	0.95	-1.76	1.90	4.51	0.64	3.56	4.07
1.35	0.45	3.16	-0.15	0.54	-1.64	2.29	4.53	0.74	3.19	4.30
1.50	0.65	3.49	0.01	0.32	-1.46	2.61	4.38	0.53	3.04	4.51
1.65	0.67	3.88	0.22	0.15	-1.30	3.15	4.76	0.28	2.83	4.83
1.80	0.53	4.22	0.63	0.02	-1.27	3.42	4.66	0.24	2.54	5.06
1.95	0.50	4.45	1.07	-0.30	-1.22	3.89	4.82	0.00	2.41	5.40
2.10	0.35	4.99	1.42	-0.40	-1.11	4.58	4.77	0.03	1.97	5.83
2.25	0.35	5.36	1.68	-0.90	-1.02	4.82	5.12	0.35	1.78	6.54
2.40	0.13	5.71	2.49	-1.37	0.89	5.42	5.23	0.46	1.53	6.68
2.55	0.39	6.51	2.57	-1.65	0.89	6.07	5.40	0.88	1.04	7.36
2.70	0.14	7.35	3.09	-2.00	-1.02	6.44	5.84	1.27	0.86	7.91
2.85	0.14	8.02	3.40	-2.42	0.97	7.15	5.86	1.68	0.48	8.39
3.00	0.09	8.96	4.00	-3.13	0.99	7.66	6.01	1.98	-0.09	8.98

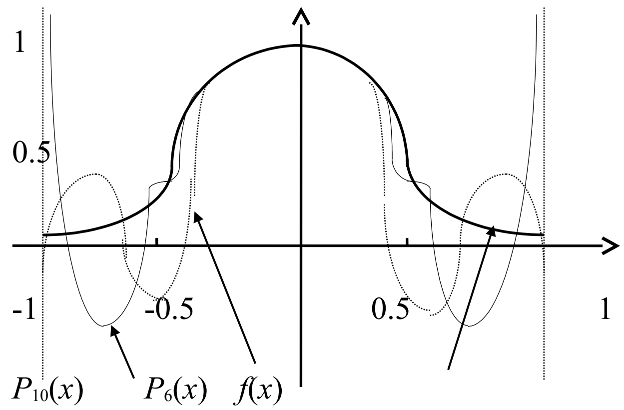
3.3. Глобальная полиномиальная интерполяция

Пусть функция $y = f(x)$ интерполируется на отрезке $[a, b]$. **Метод решения этой задачи единым для всего отрезка многочленом $P_n(x)$ называется глобальной полиномиальной интерполяцией**. Надежда приблизить $y = f(x)$ везде на $[a, b]$ с заданной точностью ε единым многочленом $P_n(x)$ базируются на теореме Вейерштрасса*.

Теорема 3.1. **Пусть функция $y = f(x)$ непрерывна на $[a, b]$. Тогда для любого $\varepsilon > 0$ существует полином $P_n(x)$ степени $n = n(\varepsilon)$ такой, что $\max\limits_{[a,b]}|f(x) - P_n(x)| < \varepsilon$.**

Однако существуют очень веские причины, по которым глобальная интерполяция многочленами высокой степени в вычислительной практике не используется. Обычный подход увеличения точности интерполяции - увеличение числа узлов. Однако не существует единой для всех непрерывных на отрезке $[a, b]$ функций стратегии выбора узлов интерполяции. Чаще всего узлы располагаются на $[a, b]$ равномерно. Но даже для очень гладких функций это иногда не приносит желаемого эффекта.

* Карл Вейерштрасс (1815-1897) – немецкий математик.



Классический пример - функция Рунге* $f(x)=\dfrac{1}{1+25x^2}$, $x\in[-1,1]$. Если использовать глобальную аппроксимацию на $[-1,1]$ с равномерным распределением узлов, то при больших n интерполяция дает очень хорошие результаты в центральной части отрезка. В то же время последовательность $P_n(x)$ расходится при $n\to\infty$ для $0.73<|x|\le 1$ (см. рисунок слева). Таким образом, равномерное распределение узлов оказалось неудачным. Итак, не существует единой для всякой функции $y=f(x)$ стратегии выбора узлов интерполяции. Об этом же говорит и теорема Фабера*.

Теорема 3.2. **Какова бы ни была стратегия выбора узлов интерполяции, найдется непрерывная на отрезке $[a,b]$ функция $y=f(x)$, для которой** $\max\limits_{[a,b]}|f(x)-P_n(x)|\to\infty$ при $n\to\infty$.

Однако если функция гладкая (непрерывно дифференцируемая), то такая стратегия существует.

Теорема 3.3. **Пусть в качестве узлов интерполяции на отрезке $[a,b]$ выбираются узлы полиномов Чебышева*. Тогда для любой непрерывно дифференцируемой на отрезке $[a,b]$ функции $y=f(x)$ метод интерполяции сходится.**

Практическая реализация стратегии выбора узлов интерполяции возможна и оправдана в довольно редких случаях и просто невозможна тогда, когда приходится иметь дело с заданной таблицей значений функции.

3.4. Чувствительность интерполяционного многочлена к погрешностям входных данных

Помимо погрешности от приближенной замены $y=f(x)$ на $P_n(x)$ возникает еще дополнительная погрешность, связанная с тем, что значения интерполируемой функции тоже задаются с погрешностью. Пусть заданные в узлах x_i, $i=\overline{0,n}$ значения y_i^* содержат погрешности ε_i. Тогда $P_n^*(x)=\sum\limits_{j=0}^{n} y_j^* l_{nj}(x)$ содержат погрешность $P_n(x)-P_n^*(x)=\sum\limits_{j=0}^{n}\varepsilon_j l_{nj}(x)$, где $\{l_{nj}(x)\}_{j=0}^{n}$ - Лагранжев базис.

Пусть известно, что верхняя граница погрешности y_i^* равна $\overline{\Delta}(y^*)$, то есть $|\varepsilon_i|\le\overline{\Delta}(y^*)$ для $\forall i=\overline{0,n}$. Тогда для верхней границы соответствующей погрешности многочлена $\overline{\Delta}(P_n^*)=\max\limits_{[a,b]}|P_n(x)-P_n^*(x)|$ справедлива оценка

$$\overline{\Delta}(P_n^*(x))\le\Lambda_n\overline{\Delta}(y^*),\ \text{где}$$

$$\Lambda_n=\max_{[a,b]}\sum_{j=0}^{n}|l_{nj}(x)|\ \text{- константа Лебега}^{*}. \qquad (3.4.1)$$

В задаче интерполирования константа Лебега играет роль абсолютного числа обусловленности, то есть в самом неблагоприятном случае погрешность входных данных

** Карл Давид Тольме Рунге (1856-1927) - немецкий физик и математик.

* Жорж Фабер (1877-1966) - швейцарский математик.

** Пафнутий Львович Чебышев (1821-1894) – русский математик и механик.

*** Анри Леон Лебег (1875-1941) - французский математик.

при интерполяции может возрасти в Λ_n раз. Величина Λ_n зависит от расположения узлов интерполяции. Например, если в качестве узлов интерполяции взяты нули многочленов Чебышева, то

$$\Lambda_n \approx \frac{2}{\pi}\ln(n+1)+1. \qquad (3.4.2)$$

Если же узлы равноотстоящие, то $\Lambda_n > \frac{2^{n-1}}{(2n-1)\sqrt{n}}$ и уже при $n \geq 4$ обусловленность задачи резко ухудшается. Из этого следует важный практический вывод: в вычислениях_не следует использовать интерполяционные многочлены высокой степени с равноотстоящими узлами.

3.5. Многочлены Чебышева

Система функций $\{\varphi_n(x)\}$, заданных на $[a,b]$, называется ортогональной на $[a,b]$, если

$$\begin{aligned}
&1.\ \int_a^b \varphi_n^2(x)dx = d_n \neq 0 \text{ и } \exists \text{ для } \forall n = 0,1,2,...\\
&2.\ \int_a^b \varphi_n(x)\varphi_m(x)dx = \begin{cases} 0, & m \neq n,\\ d_n \neq 0, & n = m.\end{cases}
\end{aligned} \qquad (3.5.1)$$

Система функций $\{\varphi_n(x)\}$, заданных на $[a,b]$, называется ортогональной на $[a,b]$ с весом $\rho(x)$, если

$$\begin{aligned}
&1.\ \int_a^b \varphi_n^2(x)\rho(x)dx = d_n \neq 0 \text{ и } \exists \text{ для } \forall n = 0,1,2,...\\
&2.\ \int_a^b \varphi_n(x)\varphi_m(x)\rho(x)dx = \begin{cases} 0, & m \neq n,\\ d_n \neq 0, & n = m.\end{cases}
\end{aligned} \qquad (3.5.2)$$

Функция $\rho(x)$ называется весовой функцией для системы $\{\varphi_n(x)\}$. Если $d_n = 1$ для $\forall n$, то система функций $\{\varphi_n(x)\}$ называется ортонормированной.

В качестве примера системы функций, ортогональной с весом, приведем многочлены Чебышева, которые известны еще и тем, что являются полиномами, наименее уклоняющимися от нуля. Эти многочлены определяют разными способами. Например:

1. $T_n(x) = \frac{1}{2}\left[\left(x+\sqrt{x^2-1}\right)^n + \left(x-\sqrt{x^2-1}\right)^n\right],\ n = 1,2,3,...$ (3.5.3)

2. Являются решениями следующего дифференциального уравнения:

$$\left(1-x^2\right)T_n^{//}(x) - xT_n^{/}(x) + n^2T_n(x) = 0. \qquad (3.5.4)$$

3. Определяются из формулы Родрига*

$$\frac{T_n(x)}{\sqrt{1-x^2}} = \frac{(-1)^n}{n!}\frac{d^n\left[\left(1-x^2\right)^{n-\frac{1}{2}}\right]}{dx^n},\ n = 0,1,2,... \qquad (3.5.5)$$

4. Определяются рекуррентно:

$$\begin{aligned}
&T_0(x) = 1,\ T_1(x) = x,\\
&T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x),\ n = 1,2,3,...
\end{aligned} \qquad (3.5.6)$$

Иногда в качестве полиномов Чебышева берут функции $\widetilde{T}_n(x) = \frac{1}{2^n}T_n(x)$.

5. $T_n(x) = \cos(n \cdot \arccos x),\ n = 0,1,2,...\ x \in [-1, 1]$. (3.5.7)

* Родриг; Бенжамен Оленд Родригес (1794-1851) - французский математик и экономист.

6. $T_n(x)=\dfrac{n}{2}\sum\limits_{m=0}^{\left[\frac{n}{2}\right]}(-1)^m\dfrac{(n-m-1)!}{m!(n-2m)!}(2x)^{n-2m}.$ (3.5.8)

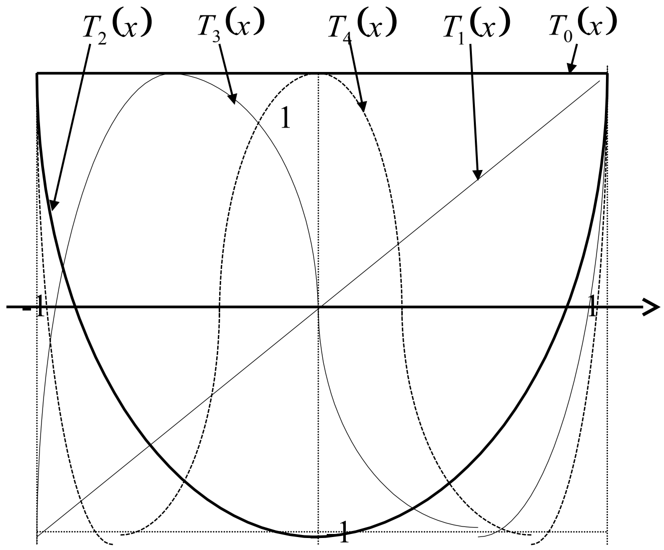
Многочлены Чебышева обладают множеством замечательных свойств.

Теорема 3.4. **Полиномы Чебышева $T_n(x)$ образуют на отрезке $[-1,1]$ ортогональную систему с весом $\rho(x)=\frac{1}{\sqrt{1-x^2}}$, то есть**

$$\int\limits_{-1}^{1}T_n(x)T_m(x)\frac{dx}{\sqrt{1-x^2}}=\begin{cases}0, & n\neq m,\\ \pi, & n=m=0,\\ \frac{\pi}{2}, & n=m>0.\end{cases} \qquad (3.5.9)$$

Действительно, $\int\limits_{-1}^{1}T_n(x)T_m(x)\frac{dx}{\sqrt{1-x^2}}=\left\langle\begin{array}{c}x=\cos t,\\ 0\le t\le\pi,\\ T_n(\cos t)=\cos(n\cdot\arccos(\cos t))=\cos nt\end{array}\right\rangle=$

$$=\int\limits_{\pi}^{0}\cos nt\cdot\cos mt\frac{(-\sin t)}{|\sin t|}dt=\frac{1}{2}\int\limits_{-\pi}^{\pi}\cos nt\cdot\cos mt\,dt=\begin{cases}0, & n\neq m,\\ \pi, & n=m,\\ \frac{\pi}{2}, & n=m>0.\end{cases}.$$



$T_0(x)=1,$
$T_1(x)=x,$
$T_2(x)=2x^2-1,$
$T_3(x)=4x^3-3x,$
$T_4(x)=8x^4-8x^2+1,$
$T_5(x)=16x^5-20x^3+5x$
и так далее.

Теорема 3.5. **При четном n многочлен $T_n(x)$ содержит только четные степени x и является четной функцией, а при нечетном n многочлен $T_n(x)$ содержит только нечетные степени x и является нечетной функцией.**

Теорема 3.6. **При $n\ge 1$ старший коэффициент многочлена $T_n(x)$ равен 2^{n-1}, то есть $T_n(x)=2^{n-1}x^n+\dots$**

Теорема 3.7. **При $n\ge 1$ многочлен $T_n(x)$ имеет ровно n действительных корней, расположенных на отрезке $[-1,1]$ и вычисляемых по формуле**

$$x_k=\cos\frac{(2k+1)\pi}{2n},\ k=0,1,2,\dots,n-1. \qquad (3.5.10)$$

Теорема 3.8. **При $n\ge 0$ справедливо равенство $\max\limits_{[-1,1]}|T_n(x)|=1$. Если $n\ge 1$, то этот максимум достигается ровно в $n+1$ точках, которые находятся по формуле**

$$x_m=\cos\frac{\pi m}{n},\ m=0,1,\dots,n. \qquad (3.5.11)$$

При этом $T_n(x_m) = (-1)^m$, то есть максимумы и минимумы многочлена Чебышева чередуются.

Теоремы 3.7 и 3.8 легко доказываются с помощью формулы (3.5.7).

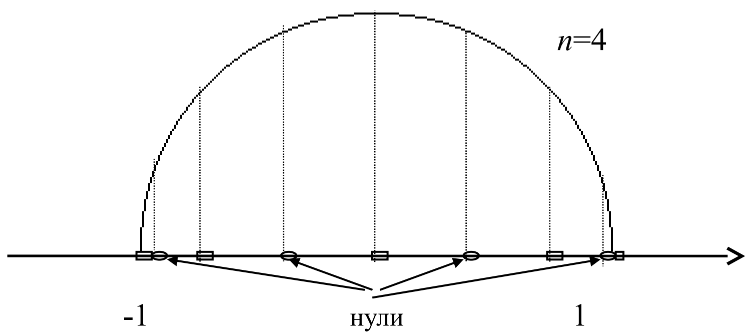
Назовем величину $\max\limits_{[-1,1]}|P_n(x)|$ уклонением многочлена $T_n(x)$ от нуля. Тогда справедлива следующая, доказанная П.Л. Чебышевым в 1854 г. теорема.

Теорема 3.9. **Среди всех многочленов фиксированной степени $n \geq 1$ со старшим коэффициентом a_n, равным единице, наименьшее уклонение от нуля, равное 2^{1-n}, имеет многочлен $\widetilde{T}_n(x) = \frac{1}{2^{n-1}} T_n(x)$.**

Последнее свойство имеет особую ценность для приложений. Действительно, тогда для любого многочлена $P_n(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0$, отличного от $\widetilde{T}_n(x)$, справедливо неравенство

$$\frac{1}{2^{n-1}} = \max_{[-1,1]}\left|\widetilde{T}_n(x)\right| < \max_{[-1,1]}|P_n(x)|. \qquad (3.5.12)$$

В силу формул (3.5.10) и (3.5.11) нули и точки экстремума полинома $T_n(x)$ можно построить следующим образом: разделив полуокружность, опирающуюся на отрезок $[-1,1]$ на $2n$ частей, спроецируем полученные точки на диаметр. Нумеруя проекции слева направо, получим, что все проекции с нечетными номерами являются нулями полинома $T_n(x)$, а с четными – его точками экстремума. Из геометрических соображений вытекает, что как нули, так и точки экстремума полинома $T_n(x)$ сгущаются к концам отрезка $[-1,1]$.



3.6. Решение задачи минимизации оценки погрешности

Предположим, что значение заданной на отрезке $[a,b]$ функции $f(x)$ можно вычислить в произвольной точке x. Однако по некоторым причинам целесообразнее заменить прямое вычисление функции $f(x)$ вычислением значений ее интерполяционного многочлена. Для такой замены необходимо один раз получить таблицу значений $f(x)$ в выбранных точках $x_0, x_1, \ldots, x_n \in [a,b]$. При этом естественно стремиться к такому выбору узлов интерполяции, который позволит сделать минимальной величину $\Delta(P_n(x)) = \max\limits_{[a,b]}|f(x) - P_n(x)|$.

Пусть о функции $f(x)$ известно лишь то, что она непрерывно дифференцируема $n+1$ раз на отрезке $[a,b]$. Тогда

$$\overline{\Delta}(P_n(x)) = \frac{M_{n+1}}{(n+1)!} \max_{[a,b]}|\omega_{n+1}(x)|. \qquad (3.6.1)$$

Найдем теперь набор узлов интерполяции $x_0, x_1, \ldots, x_n$, при котором $\overline{\Delta}(P_n(x)) \to \min$. Пусть сначала $[a,b] = [-1,1]$. В этом случае величина (3.6.1) будет минимальна, если будет минимальна $\max\limits_{[-1,1]}|\omega_{n+1}(x)|$. Но этим свойством обладают полиномы Чебышева, следовательно, $\omega_{n+1}(x) \equiv \widetilde{T}_{n+1}(x)$, и набор узлов определен $x_k = \cos\dfrac{2k+1}{2(n+1)}\pi,\ k = 0,1,...,n$. Это нули многочлена $\widetilde{T}_{n+1}(x)$. При таком выборе

$$\overline{\Delta}(P_n(x)) = \frac{M_{n+1}}{2^n(n+1)!}, \qquad (3.6.2)$$

причем любой другой выбор узлов дает большее значение верхней границы погрешности. Для формулы Тейлора, например, $\overline{\Delta}(P_n(x)) = \dfrac{M_{n+1}}{(n+1)!}$, то есть в 2^n раз хуже.

Перейдем теперь к отрезку $[a,b]$. Его можно преобразовать к стандартному отрезку $[-1,1]$ следующей заменой: $x = \dfrac{a+b}{2} + \dfrac{b-a}{2}t$, где $t \in [-1,1]$. В этом случае

$$\omega_{n+1}(x) = \left[\frac{b-a}{2}\right]^{n+1} \widetilde{\omega}_{n+1}(t),$$
$$\text{где } \widetilde{\omega}_{n+1}(t) = (t-t_0)(t-t_1)...(t-t_n)$$
$$\text{и } x_k = \frac{a+b}{2} + \frac{b-a}{2}t_k,\ k = 0,1,...,n. \qquad (3.6.3)$$

Тогда
$$\overline{\Delta}(P_n(x)) = \frac{M_{n+1}}{(n+1)!}\left[\frac{b-a}{2}\right]^{n+1} \max_{[-1,1]}|\widetilde{\omega}_{n+1}(t)| \qquad (3.6.4)$$

и минимум этой величины достигается при значениях $t_0, t_1, ..., t_n$, совпадающих с нулями многочлена $\widetilde{T}_{n+1}(x)$. Таким образом, решение поставленной задачи дает выбор узлов

$$x_k = \frac{a+b}{2} + \frac{b-a}{2}\cos\left(\frac{2k+1}{2n+2}\pi\right),\ k = 0,1,...,n, \qquad (3.6.5)$$

и ему соответствует минимальное значение верхней границы погрешности интерполяции, равное:
$$\overline{\Delta}(P_n(x)) = \frac{M_{n+1}}{2^n(n+1)!}\left[\frac{b-a}{2}\right]^{n+1}. \qquad (3.6.6)$$

3.7. Ряд Фурье* по многочленам Чебышева

Функциональный ряд вида $f(x) = \sum\limits_{i=0}^{\infty} c_i \varphi_i(x)$, где $\{\varphi_i(x)\}$ **- система базисных функций, называется рядом Фурье функции** $f(x)$ **на отрезке** $[a,b]$ **по системе функций** $\varphi_i(x),\ i = 0,1,...$ **с весом** $\rho(x)$ **на** $[a,b]$**, если коэффициенты** c_i **вычисляются по формулам вида**

$$c_i = \frac{1}{\int\limits_a^b \varphi_i^2(x)\rho(x)dx}\int\limits_a^b f(x)\varphi_i(x)\rho(x)dx,\ i = 0,1,...$$

(3.7.1)

Так как для полиномов Чебышева весовая функция равна: $\rho(x) = \dfrac{1}{\sqrt{1-x^2}}$ на $[-1,1]$,

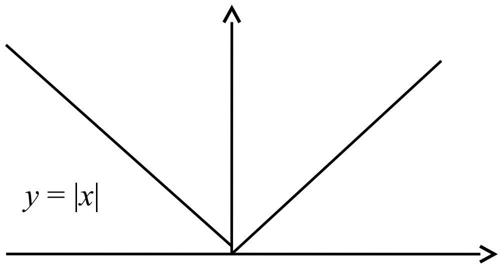
то
$$\begin{cases} f(x) = \sum\limits_{i=0}^{\infty} a_i T_i(x), \text{ где} \\ a_0 = \dfrac{1}{\pi}\int\limits_{-1}^{1}\dfrac{f(x)}{\sqrt{1-x^2}}dx,\ a_i = \dfrac{2}{\pi}\int\limits_{-1}^{1}\dfrac{f(x)T_i(x)}{\sqrt{1-x^2}}dx,\ i = 1,2,... \end{cases}$$

(3.7.2)

Пример. Разложить в ряд Фурье по полиномам Чебышева функцию $y = |x|$ на отрезке $[-1,1]$.

* Жан Батист Жозеф Фурье (1768-1830) - французский математик.

Функция $y = |x|$ - четная, поэтому, как во всяком ряду Фурье, нечетные коэффициенты будут равны нулю, четные же можно сдвоить, уменьшив при этом интервал интегрирования вдвое. Тогда $a_0 = \frac{2}{\pi}\int_0^1 \frac{f(x)}{\sqrt{1-x^2}}dx,\ a_i = \frac{4}{\pi}\int_0^1 \frac{f(x)T_i(x)}{\sqrt{1-x^2}}dx.$ Вычислим коэффициенты разложения:



$$a_0 = \frac{2}{\pi}\int_0^1 \frac{xdx}{\sqrt{1-x^2}} = \left\langle \begin{array}{c} x = \cos t,\ dx = -\sin tdt, \\ \sqrt{1-x^2} = |\sin t|,\ x = 0, t = \frac{\pi}{2}, \\ x = 1, t = 0. \end{array} \right\rangle =$$

$$= \frac{2}{\pi}\int_0^{\frac{\pi}{2}} \cos tdt = \frac{2}{\pi}\sin t\Big|_0^{\frac{\pi}{2}} = \frac{2}{\pi}.$$

$$a_i = \frac{4}{\pi}\int_0^1 \frac{xT_i(x)dx}{\sqrt{1-x^2}} =$$

$$-\frac{4}{\pi}\int_{\frac{\pi}{2}}^0 \frac{\cos t\sin t\cos it}{\sin t}dt = \frac{4}{\pi}\int_0^{\frac{\pi}{2}} \cos t\cos itdt =$$

$$= \frac{2}{\pi}\int_0^{\frac{\pi}{2}} [\cos(t(i-1)) + \cos(t(i+1))]dt = \frac{2}{\pi}\left[\frac{1}{i-1}\sin(t(i-1))\Big|_0^{\frac{\pi}{2}} + \frac{1}{i+1}\sin(t(i+1))\Big|_0^{\frac{\pi}{2}}\right] =$$

$$= \frac{2}{\pi}\left[\frac{-1}{i-1}\cos\frac{\pi i}{2} + \frac{1}{i+1}\cos\frac{\pi i}{2}\right] = -\frac{4}{\pi(i^2-1)}\cos\frac{\pi i}{2} = \begin{cases} 0,\ i = 2m+1, \\ \frac{4}{\pi(4m^2-1)}(-1)^{m+1},\ i = 2m. \end{cases}$$

Итак, $|x| = \frac{2}{\pi} + \frac{4}{\pi}\sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{(4m^2-1)} T_{2m}(x).$

Пример. По равномерной (а) и специальной (б) таблице значений функции $y = \lg x$ найти $\lg 0.5$ и оценить погрешность. Использовать формулы Ньютона с разделенными разностями.

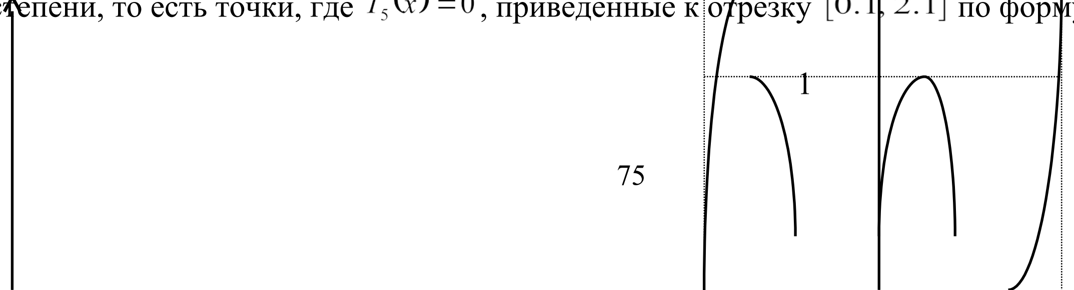
(а)

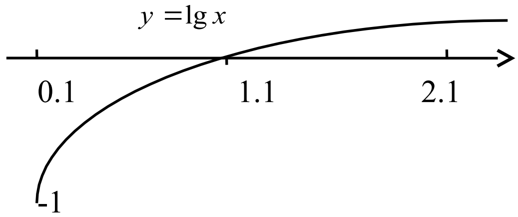
x	0.1	0.6	1.1	1.6	2.1
y	-1.000000	-0.221849	0.041393	0.204120	0.322219

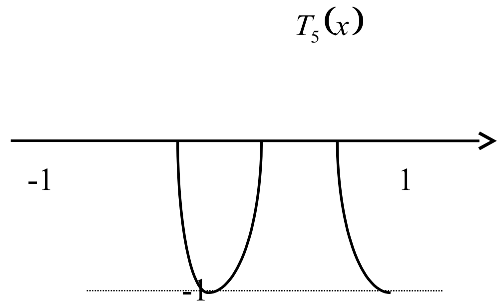
(б)

x	0.148944	0.512215	1.1	1.687785	2.051057
y	-0.826977	-0.290548	0.041393	0.227317	0.311978

Так как таблицы короткие, то по ним можно вычислить разделенные разности лишь до четвертого порядка. Точки аргумента в таблице (б) - это нули многочлена Чебышева пятой степени, то есть точки, где $T_5(x) = 0$, приведенные к отрезку $[0.1, 2.1]$ по формуле







$x_k = \frac{a+b}{2} + \frac{b-a}{2}\cos\left(\frac{2k+1}{2n+2}\right),\ k = 0,1,2,3,4,\ n = 4.$ Действительно, $a = 0.1,\ b = 2.1,$

$$k = 0,\ x_0 = 1.1 + \cos\frac{\pi}{10} = 2.051057,$$
$$k = 1,\ x_1 = 1.1 + \cos\frac{3\pi}{10} = 1.687785,$$
$$k = 2,\ x_2 = 1.1 + \cos\frac{\pi}{2} = 1.100000,$$
$$k = 3,\ x_3 = 1.1 + \cos\frac{7\pi}{10} = 0.512215,$$
$$k = 4,\ x_4 = 1.1 + \cos\frac{9\pi}{10} = 0.148944.$$

Согласно теории, в случае расположения узлов интерполяции в нулях многочлена Чебышева гарантирована минимальная погрешность интерполяции, равная $\overline{\Delta}(P_n(x)) = \frac{M_{n+1}}{2^n(n+1)!}\left[\frac{b-a}{2}\right]^{n+1}$, в отличие от обычной интерполяции по формуле Ньютона с разделенными разностями, где погрешность равна: $\overline{\Delta}(P_n(x)) = \frac{M_{n+1}}{(n+1)!}\max\limits_{[a,b]}|\omega_{n+1}(x)|.$

Построим для случаев (а) и (б) таблицы разделенных разностей.

(а)

i	x	y	$f(x_0;x_1)$	$f(x_0;x_1;x_2$	$f(x_0;x_1;x_2;x$	$f(x_0;x_1;\ldots;$
0	0.1	-1.000000				
			1.556302			
1	0.6	-0.221849		-1.029818		
			0.526484		0.552525	
2	1.1	0.041393		-0.201030		-0.239005
			0.325454		0.074516	
3	1.6	0.204120		-0.089256		
			0.236198			
4	2.1	0.322219				

(б)

i	x	y	$f(x_0;x_1)$	$f(x_0;x_1;x_2$	$f(x_0;x_1;x_2;$	$f(x_0;x_1;\ldots$
0	0.148944	-0.826977				
			1.476663			
1	0.512215	-0.290548		-0.958862		
			0.564732		0.485784	
2	1.100000	0.041393		-0.211318		-0.213107
			0.316313		0.080431	
3	1.687785	0.227317		-0.087547		
			0.233051			
4	2.051057	0.311978				

Вычисления по этим таблицам дают:

(а)

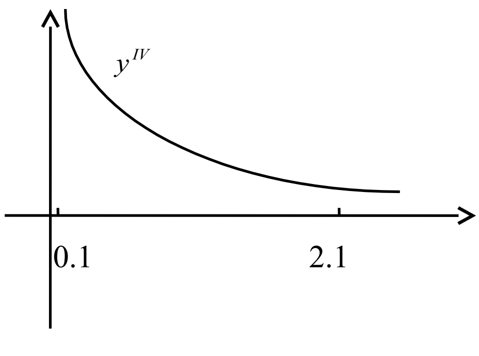
$$\lg 0.5 = -1.000000 + 1.556302 \cdot 0.4 - 1.029818 \cdot 0.4 \cdot (-0.1) + 0.552525 \cdot 0.4 \cdot (-0.1) \cdot (-0.6) -$$
$$- 0.239005 \cdot 0.4 \cdot (-0.1) \cdot (-0.6) \cdot (-1.1) = -1.000000 + 0.622521 + 0.041193 + 0.013261 +$$
$$+ 0.006310 = -0.316716.$$

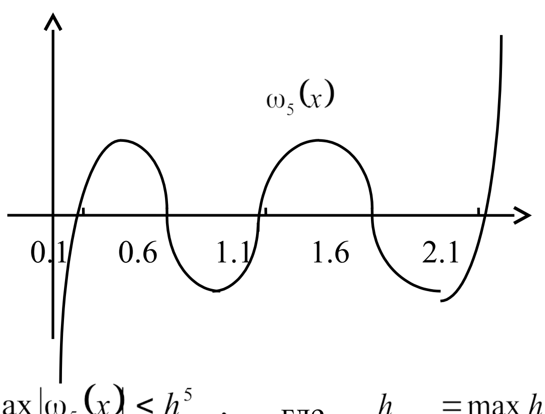
(б)

$$\lg 0.5 = -0.826977 + 1.476663 \cdot 0.351056 - 0.958862 \cdot 0.351056 \cdot (-0.012215) + 0.485784 \cdot$$
$$\cdot 0.351056 \cdot (-0.012215) \cdot (-0.6) - 0.213107 \cdot 0.351056 \cdot (-0.012215) \cdot (-0.6) \cdot (-1.187785) =$$
$$= -0.826977 + 0.518391 + 0.004112 + 0.001250 + 0.000651 = -0.302573.$$

Истинное значение $\lg 0.5 = -0.301030$, таким образом, предельная погрешность в случае (а) достигает $\bar{\Delta}_а = 0.01569$, а в случае (б) $\bar{\Delta}_б = 0.00154$, то есть на порядок меньше.

Так как таблица разностей очень короткая, а $y = \lg x$ вблизи точки $x = 0.1$ меняется быстро (см. график функции $y = \lg x$), то попытка оценить погрешность по известным формулам не даст достоверного результата, то есть погрешность будет слишком завышена. Действительно, $y = \lg x,\ y^{\mathrm{IV}} = \dfrac{24}{x^5} \lg e,\ \max\limits_{[0.1,\,2.1]} \left|y^{\mathrm{IV}}\right| > 1 \cdot 10^6.$ Так же, огрубляя оценку и вспоминая график





функции $\omega_5(x)$, будем иметь $\max\limits_{[0.1,\,2.1]} |\omega_5(x)| < h_{\max}^5$, где $h_{\max} = \max\limits_{0 \le i \le 4} h_i$. Тогда $R_а \le \dfrac{1 \cdot 10^6 \cdot 0.03125}{120} = 260.$ Аналогично для случая (б) $R_б \le \dfrac{1.4 \cdot 10^5 \cdot 1.61051}{120 \cdot 32} = 58.$ Оба результата не реальны.