УДК 517 **Физико-математические науки**

Тюнькин Андрей Борисович, *студент 2 курс, нефтегазовый факультет, ФГБОУ «Ухтинский государственный технический университет»*

Басов Александр Андреевич, *студент 2 курс, нефтегазовый факультет, ФГБОУ «Ухтинский государственный технический университет»*

Пармузина Мария Семеновна, *научный руководитель, доцент кафедры высшей математики Ухтинский государственный технический университет, кандидат педагогических наук*

Россия, г. Ухта

АППРОКСИМАЦИЯ С ПОМОЩЬЮ СРЕДНЕКВАДРАТИЧЕСКОГО ПРИБЛИЖЕНИЯ

Аннотация: При проведении инженерных исследований часто возникает задача аппроксимации – составление функции по отдельным значениям. Значения могут быть получены из эксперимента, из наблюдений или из каких-то расчетов (например, при решении дифференциального уравнения численными методами, получается таблица соответствия значений переменной и приближенных значений решения). Знание основных методов и формул составления аппроксимирующих функций является профессионально значимым для студентов технического вуза, поэтому в своей работе мы рассмотрели основные идеи аппроксимации, изучили теоретические сведения и выполнили некоторые практические расчеты.

Ключевые слова: аппроксимация, сглаживание, среднеквадратическое приближение, эмпирическая формула.

Annotation: When conducting engineering research, the problem of approximation often arises – composing a function based on individual values. The values can be obtained from an experiment, from observations, or from some

calculations (for example, when solving a differential equation by numerical methods, a table of correspondence between the values of the variable and the approximate values of the solution is obtained). Knowledge of the basic methods and formulas for drawing up approximating functions is professionally important for students of a technical university, so in our work we considered the main ideas of approximation, studied theoretical information and performed some practical calculations.

Keywords: approximation, smoothing, root-mean-square approximation, empirical formula.

Часто в практической работе необходимо найти функциональную зависимость $y = \varphi(x)$ между величинами x и y, которые заданы отдельными парами значений (x_i, y_i). Эти данные могут быть получены в результате измерений, в каких-то теоретических методах или расчетах.

Многие численные методы решения математических задач основаны на замене одной функции (которая по каким-то критериям не удобна для решения задачи) на другую функцию (которая обладает хорошими качествами). Часто исходную «сложную» функцию заменяют линейной функцией, так как это наиболее простая и удобная для вычислений функция.

Задача составления функции по отдельным значениям или замена «сложной» функции на «удобную» называют аппроксимацией или приближением функции.

Рассмотрим задачу аппроксимации по отдельным значениям. Для получения решения задачи аппроксимации при этом подходе необходимо выполнение некоторой последовательности действий: надо определить общий вид приближенной функции, включающий неизвестные коэффициенты; определить конкретные значения параметров на основе заданного критерия близости исходных значений и значений аппроксимирующей функции.

В математике существует два основных подхода к определению близости в данной задаче – это интерполяция и сглаживание.

Интерполяция строится с помощью интерполяционного многочлена, который проходит непосредственно через все точки данных. В этом случае аппроксимирующая функция представляется в виде: интерполяционного многочлена Лагранжа или интерполяционного многочлена Ньютона.

Сглаживание производится с помощью построения функции, которая проходит близко от точек из заданного массива данных. Таким образом, аппроксимирующая функция сглаживает все случайные помехи (или погрешности). Проведенный анализ позволяет заключить, что аппроксимацию сглаживанием стоит применять, если исходные данные содержат погрешности; исходные данные содержат повторы; большое количество исходных данных; имеются теоретические сведения о виде аппроксимирующей функции.

Полученная сглаживанием функция $y = \varphi(x)$ называется эмпирической формулой. Эмпирические формулы это не законы природы, а всего лишь гипотезы. Однако значение их весьма велико.

Задача ставится следующим образом. Даны n пар значений аргумента и функции: x_i, y_i ($i = 1, \ldots, n$). Кроме того, из теории известен общий вид функции $y = \varphi(x, a, b, c, \ldots)$, которая связывает исследуемые переменные величины x и y. Функция содержит конечное число неизвестных постоянных параметров $a, b, c,$ … . Требуется определить эти численные значения. Для вычисления параметров функции $a, b, c, \ldots$ должны использоваться экспериментальные данные x_i , y_i (i = 1, … , n).

В качестве возможной оценки качества аппроксимации можно взять максимальное значение из модулей разностей $\max\limits_{i=1,\ldots,n}\left|\varphi(x_i, a, b, c, \ldots) - y_i\right|$. Можно в качестве критерия «наилучшего» приближения выбрать среднее арифметическое абсолютных значений отклонений $\dfrac{\sum\limits_{i=1}^{n}\left|\varphi(x_i, a, b, c, \ldots) - y_i\right|}{n}$| или среднеквадратичное отклонение.

Наиболее часто применяемый метод – это среднеквадратическое отклонение. Разработка этого метода связана с именами известных математиков прошлого – К. Гаусса и А. Лежандра.

Метод наименьших квадратов – математический метод, основанный на определении аппроксимирующей функции, которая строится в ближайшей близости от точек из заданного массива экспериментальных данных. Близость исходной и аппроксимирующей функции $\varphi(x)$ определяется числовой мерой, а именно: сумма квадратов отклонений экспериментальных данных от аппроксимирующей кривой $\varphi(x)$ должна быть наименьшей.

Данный критерий метода наименьших квадратов записывается в виде следующего выражения: $\sum_{i=1}^{n} \xi_i^2 = \sum_{i=1}^{n} \left(\varphi(x_i, a, b, c, ...) - y_i\right)^2 \to \min$.

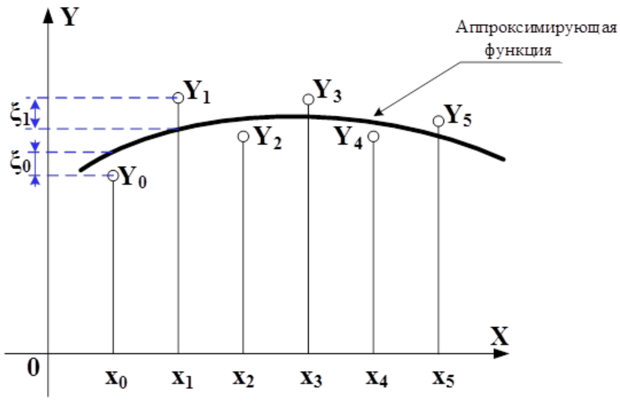


Рисунок 1. Аппроксимирующая кривая, построенная по методу наименьших квадратов

Степень точности аппроксимации исследуемого процесса с помощью полученной функциональной зависимости может быть оценена по значению среднего квадратичного отклонения.

Под средним квадратичным отклонением понимается число: $\Delta = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \varphi(x_i)\right)^2}$, где y_i – экспериментальное значение; $\varphi(x_i)$ – расчетное значение, вычисленное по аппроксимирующей формуле для x_i.

Также оценку аппроксимирующей функции дает средняя ошибка аппроксимации – среднее отклонение расчетных значений от фактических значений: $A = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \varphi(x_i)}{y_i}\right| \cdot 100\%$. Допустимый предел для использования формулы A<10%.

Часто аппроксимирующая функция представляет собой многочлен степени m: $\varphi_m(x) = a_0 + a_1 x + ... + a_{m-1} x^{m-1} + a_m x^m$.

Степень аппроксимирующей функции не зависит от числа узловых точек, но ее размерность должна быть всегда меньше размерности (количества точек) заданного массива экспериментальных данных $1 \le m \le n-1$.

В общем случае, когда требуется построить аппроксимирующий многочлен степени m для заданных табличных значений, условие минимума суммы квадратов отклонений по всем узловым точкам переписывается в следующем виде: $S = \sum_{i=1}^{n} \xi_i^2 = \sum_{i=1}^{n} \left(a_0 + a_1 x + ... + a_{m-1} x^{m-1} + a_m x^m - y_i \right)^2 \to \min$.

Необходимым условием минимума функции является равенству нулю ее частных производных по неизвестным переменным. В результате полученная система линейных алгебраических выражений будет записываться в следующем виде:

$$\begin{cases} a_0 \cdot N + a_1 \cdot \sum_{i=1}^{N} x_i + ... + a_{m-1} \cdot \sum_{i=1}^{N} x_i^{m-1} + a_m \cdot \sum_{i=1}^{N} x_i^m = \sum_{i=1}^{N} y_i \\ a_0 \cdot \sum_{i=1}^{N} x_i + a_1 \cdot \sum_{i=1}^{N} x_i^2 + ... + a_{m-1} \cdot \sum_{i=1}^{N} x_i^m + a_m \cdot \sum_{i=1}^{N} x_i^{m+1} = \sum_{i=1}^{N} y_i \cdot x_i \\ ... \\ a_0 \cdot \sum_{i=1}^{N} x_i^m + a_1 \cdot \sum_{i=1}^{N} x_i^{m+1} + ... + a_{m-1} \cdot \sum_{i=1}^{N} x_i^{2 \cdot m-1} + a_m \cdot \sum_{i=1}^{N} x_i^{2 \cdot m} = \sum_{i=1}^{N} y_i \cdot x_i^m \end{cases}$$

В результате получена система линейных уравнений размерностью m+1, которая состоит из m+1 неизвестных. Данная система может быть решена с помощью любого метода решения линейных алгебраических уравнений. В результате решения будут найдены неизвестные параметры аппроксимирующей функции, обеспечивающие минимальную сумму квадратов отклонений аппроксимирующей функции от исходных данных, т. е. наилучшее возможное квадратичное приближение

В случае если степень аппроксимирующей функции m=1, то мы аппроксимируем табличную функцию прямой линией (линейная регрессия).

Аппроксимирующая функция в этом случае будет иметь вид: $\varphi(x) = a_0 + a_1 x$.

Коэффициенты можно получить, решив систему уравнений:

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

В случае если степень аппроксимирующей функции m=2, то мы аппроксимируем табличную функцию квадратичной параболой (квадратичная аппроксимация).

Аппроксимирующая функция в этом случае будет иметь вид: $\varphi(x) = a_0 + a_1 x + a_2 x^2$.

Коэффициенты можно получить, решив систему уравнений:

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

В случае если степень аппроксимирующей функции m=3, то мы аппроксимируем табличную функцию кубической параболой (кубическая аппроксимация). Аппроксимирующая функция в этом случае будет иметь вид: $\varphi(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$.

Коэффициенты можно получить, решив систему уравнений:

$$\begin{cases} a_0 n + a_1 \sum_{i=1}^{n} x_i + a_2 \sum_{i=1}^{n} x_i^2 + a_3 \sum_{i=1}^{n} x_i^3 = \sum_{i=1}^{n} y_i, \\ a_0 \sum_{i=1}^{n} x_i + a_1 \sum_{i=1}^{n} x_i^2 + a_2 \sum_{i=1}^{n} x_i^3 + a_3 \sum_{i=1}^{n} x_i^4 = \sum_{i=1}^{n} x_i y_i, \\ a_0 \sum_{i=1}^{n} x_i^2 + a_1 \sum_{i=1}^{n} x_i^3 + a_2 \sum_{i=1}^{n} x_i^4 + a_3 \sum_{i=1}^{n} x_i^5 = \sum_{i=1}^{n} x_i^2 y_i, \\ a_0 \sum_{i=1}^{n} x_i^3 + a_1 \sum_{i=1}^{n} x_i^4 + a_2 \sum_{i=1}^{n} x_i^5 + a_3 \sum_{i=1}^{n} x_i^6 = \sum_{i=1}^{n} x_i^3 y_i, \end{cases}$$

Аппроксимация линеаризацией

Многие нелинейные функции, зависящие от двух параметров, можно линеаризовать путем замены переменных. Для этого необходимо подобрать такое преобразование исходной зависимости, в результате которого она приобретает линейный вид $Y = AX + B$. Далее решается задача линейной аппроксимации для новой зависимости, и вычисленные коэффициенты A и B пересчитываются в a и b.

№	Функция	Линеаризованная форма $Y = AX + B$	Замена переменных и констант			
			X	Y	a	b
1.	$y=\frac{a}{x}+b$	$y=a\frac{1}{x}+b$	$\frac{1}{x}$	y	A	B
2.	$y=\frac{a}{x+b}$	$y=\frac{-1}{b}(xy)+\frac{a}{b}$	xy	y	$-\frac{B}{A}$	$-\frac{1}{A}$
3.	$y=\frac{x}{ax+b}$	$\frac{1}{y}=b\frac{1}{x}+a$	$\frac{1}{x}$	$\frac{1}{y}$	B	A
4.	$y=a\ln x+b$	$y=a\ln x+b$	$\ln x$	y	A	B
5.	$y=be^{ax}$	$\ln y=ax+\ln b$	x	$\ln y$	A	e^B
6.	$y=bx^a$	$\ln y=a\ln x+\ln b$	$\ln x$	$\ln y$	A	e^B

Рисунок 2. Таблица замены переменных для метода линеаризации

Составление и визуализацию аппроксимирующих функций легко проводить с помощью программы MS Excel, которая позволяет путем не сложных операций организовать непосредственное вычисление необходимых коэффициентов аппроксимирующей функции.

Приведем пример вычисления аппроксимирующих функций по реальным данным.

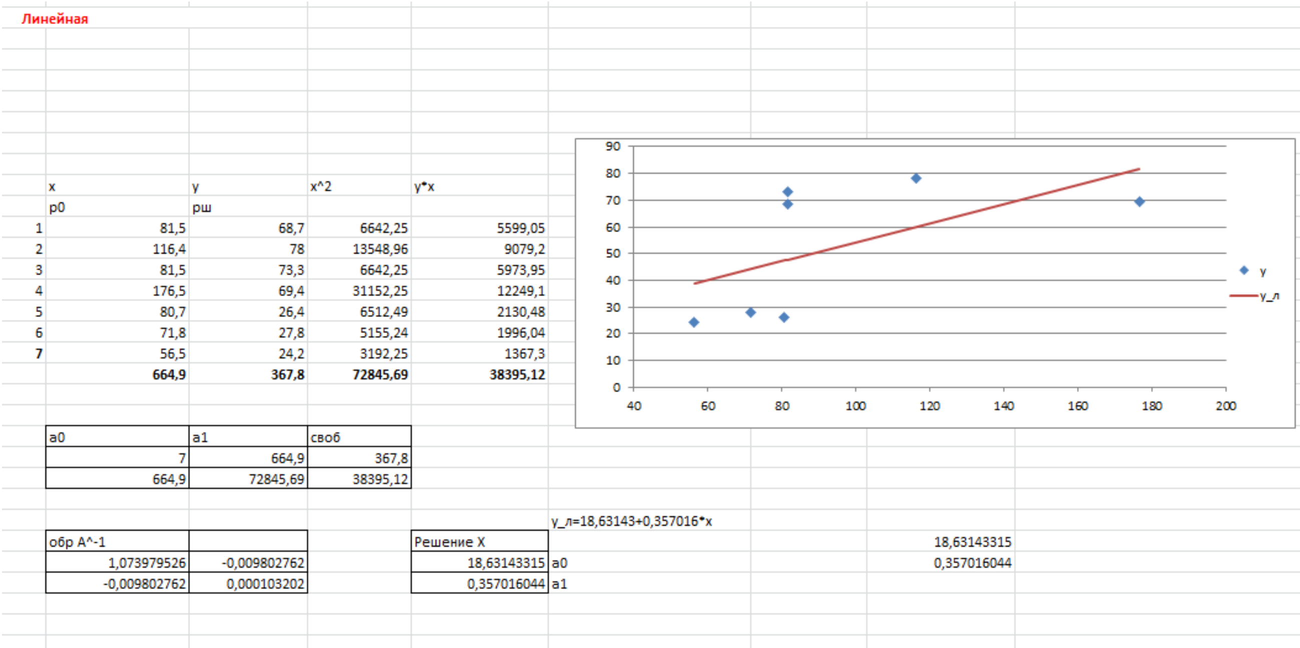
Пример. Известны результаты наблюдений за свойствами горных пород, P_0 – предел текучести по штампу; $P_ш$ – твердость по штампу.

Таблица 1.

Условие задачи

P_0	81.5	116.4	81.5	176.5	80.7	71.8	56.5
$P_ш$	68.7	78.0	73.3	69.4	26.4	27.8	24.2

Аппроксимировать линейной и квадратической функцией.



Линейная

	x	y	x^2	y*x
	p0	рш		
1	81,5	68,7	6642,25	5599,05
2	116,4	78	13548,96	9079,2
3	81,5	73,3	6642,25	5973,95
4	176,5	69,4	31152,25	12249,1
5	80,7	26,4	6512,49	2130,48
6	71,8	27,8	5155,24	1996,04
7	56,5	24,2	3192,25	1367,3
	664,9	**367,8**	**72845,69**	**38395,12**

a0	a1	своб
7	664,9	367,8
664,9	72845,69	38395,12

обр A^-1	
1,073979526	-0,009802762
-0,009802762	0,000103202

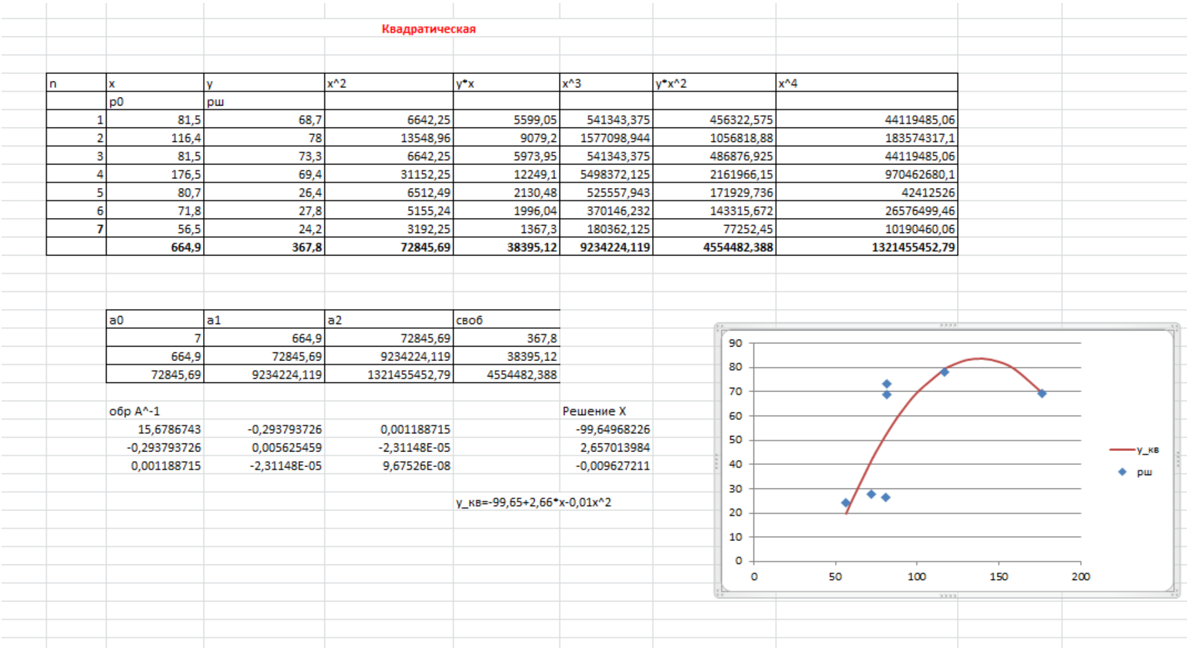
Решение X	
18,63143315	a0
0,357016044	a1

y_л=18,63143+0,357016*x

18,63143315

0,357016044

Рисунок 3. Линейная аппроксимация



Квадратическая

n	x	y	x^2	y*x	x^3	y*x^2	x^4
	p0	рш					
1	81,5	68,7	6642,25	5599,05	541343,375	456322,575	44119485,06
2	116,4	78	13548,96	9079,2	1577098,944	1056818,88	183574317,1
3	81,5	73,3	6642,25	5973,95	541343,375	486876,925	44119485,06
4	176,5	69,4	31152,25	12249,1	5498372,125	2161966,15	970462680,1
5	80,7	26,4	6512,49	2130,48	525557,943	171929,736	42412526
6	71,8	27,8	5155,24	1996,04	370146,232	143315,672	26576499,46
7	56,5	24,2	3192,25	1367,3	180362,125	77252,45	10190460,06
	664,9	**367,8**	**72845,69**	**38395,12**	**9234224,119**	**4554482,388**	**1321455452,79**

a0	a1	a2	своб
7	664,9	72845,69	367,8
664,9	72845,69	9234224,119	38395,12
72845,69	9234224,119	1321455452,79	4554482,388

обр A^-1			Решение X
15,6786743	-0,293793726	0,001188715	-99,64968226
-0,293793726	0,005625459	-2,31148E-05	2,657013984
0,001188715	-2,31148E-05	9,67526E-08	-0,009627211

y_кв=-99,65+2,66*x-0,01x^2

Рисунок 4. Квадратическая аппроксимация

Так же в программе MS Excel содержится множество встроенных функций, позволяющих в несколько секунд получить некоторые аппроксимирующие функции. Рассмотрим алгоритм получения аппроксимирующих функций.

1. Помещаем на лист таблицу с исходными данными.

2. Строим и форматируем точечную диаграмму, в которой по оси *X* задаем

значения аргумента, а по оси Y откладываем значения исходной функции, заданные таблицей.

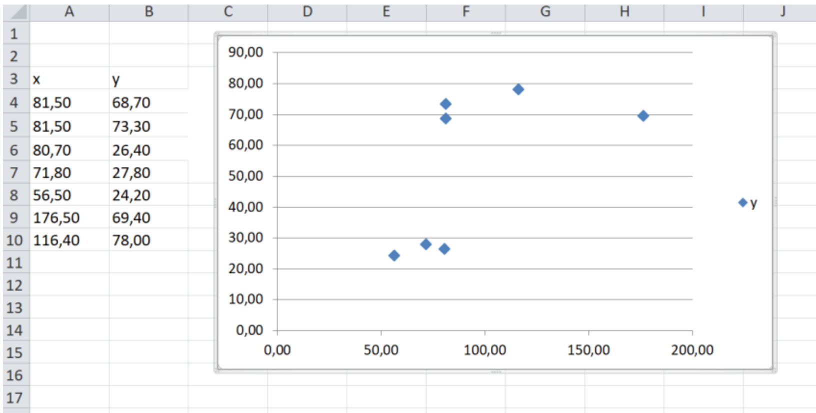


Рисунок 5. Точечная диаграмма в программе MS Excel

3. «Наводим» мышь на любую из точек на графике и щелчком правой кнопки вызываем контекстное меню и выбираем «Добавить линию тренда…». В появившемся окне выбираем тип линии и параметры отображения.

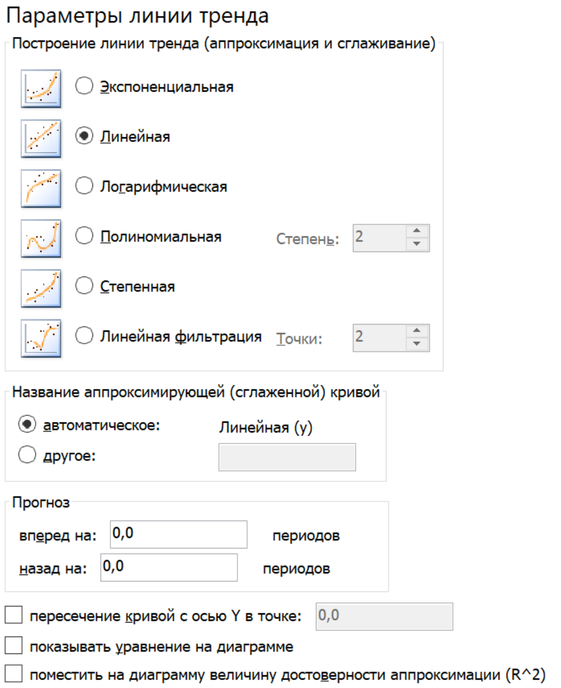


Рисунок 6. Параметры линии тренда в MS Excel

4. На графике появится выбранная линия, аппроксимирующая табличную зависимость. Кроме самой линии уравнение этой линии, можно увидеть значение

параметра R^2 – величины достоверности аппроксимации. Данный показатель может варьироваться от *0* до *1*. Чем он выше, тем аппроксимация качественнее (достовернее). Считается, что при величине данного показателя *0,85* и выше сглаживание можно считать достоверным, а если показатель ниже, то – нет.

Таким образом, проверили, что все расчеты, которые мы проводили непосредственно в MS Excel, полностью совпали с результатами полученными средствами самой программы. И этот способ является наиболее быстрым и удобным для получения стандартных аппроксимирующих функций, так как при непосредственном вычислении можно допустить ряд ошибок, которые не позволят правильно получить необходимую функцию. Но если возникнет необходимость составить аппроксимирующую функцию, которая не является стандартной в программе MS Excel, то возникнет необходимость применять непосредственный расчет коэффициентов. В этом плане, знание теоретических основ аппроксимации функции является важной составляющей для грамотного применения данного математического раздела для различных технических, инженерных расчетов.

Таким образом, проведенная работа позволила нам изучить практически важный для инженеров раздел математики – аппроксимация. Также мы вспомнили и углубили наши навыки работы с программой MS Excel, что так же является профессионально значимым навыком для грамотного специалиста.

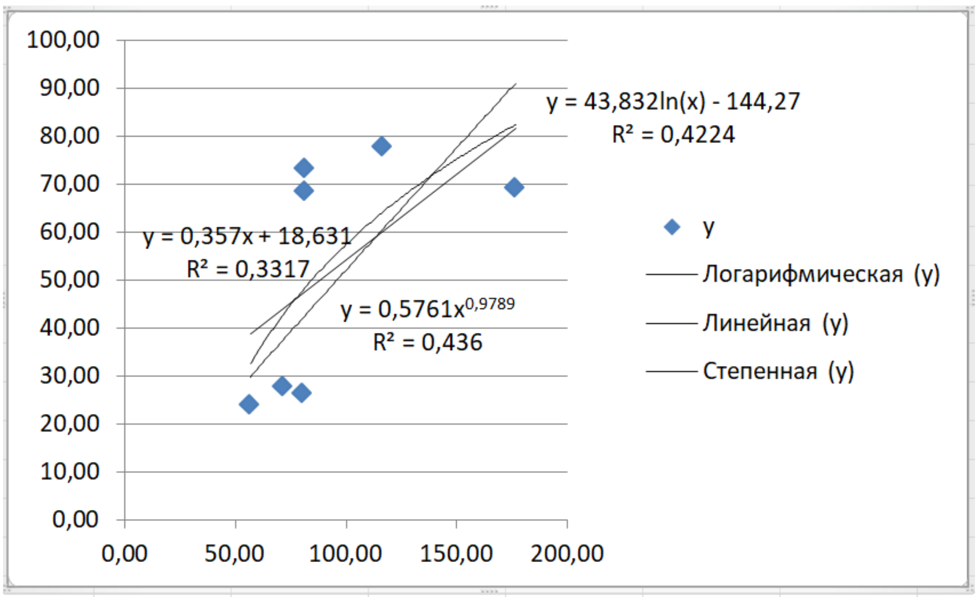


Рисунок 7. Выполненная в MS Excel аппроксимация

Библиографический список:

1. Численные методы решения задач строительства : учеб. пособие: в 2 ч. Ч. 1. / Г.Г. Кашеварова, Т.Б. Пермякова, М.Е. Лаищева. – Пермь: Изд-во Перм. нац. исслед. политехн. ун-та, 2015. – 161 с.

2. Численные методы : учебное пособие / Фаддеев М.А., Марков К.А. – Н-Новгород, ННГУ, 2010. – 158 с.