МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ»

А. Н. Гудович, Н. Н. Гудович

ЭЛЕМЕНТЫ ЧИСЛЕННЫХ МЕТОДОВ

Выпуск 3

Метод наименьших квадратов

Учебное пособие

Воронеж
Издательский дом ВГУ
2016

§ 1. Понятие о методе наименьших квадратов

Пусть требуется приблизить функцию *f* на отрезке *[a, b]* алгебраическим многочленом

$$p_n(x) = c_0 + c_1x + \ldots + c_ix^i + \ldots + c_nx^n \tag{1.1}$$

степени не выше *n*.

В качестве меры близости многочлена p_n к функции *f* в точке *x* естественно принять величину

$$| f(x) - p_n(x) |, \tag{1.2}$$

которую условимся называть *уклонением многочлена p_n от функции f в точке x*, а в качестве меры близости многочлена p_n к функции *f* на всем отрезке *[a, b]* – максимальное значение

$$\max_{a \le x \le b} | f(x) - p_n(x) | \tag{1.3}$$

уклонения (1.2) на отрезке *[a, b]*.

Величина (1.3) называется *равномерным уклонением многочлена p_n от функции f на отрезке [a, b]*. Наличие здесь наименования «равномерный» объясняется тем, что если величина (1.3) окажется меньше заданного ε, то уклонение (1.2) многочлена p_n от функции *f* в точке *x* будет меньше ε сразу для всех точек *x* из отрезка *[a, b]*, то есть будет меньше ε равномерно по *x*.

В качестве меры близости функций *f* и p_n можно также использовать среднее значение уклонений (1.2) на отрезке *[a. b]*, то есть величину

$$\frac{1}{b-a}\int_a^b |f(x) - p_n(x)|\, dx, \tag{1.4}$$

которую естественно назвать *средним уклонением p_n от f на отрезке [a, b]*.

Наконец, можно вместо уклонений (1.2) взять квадраты этих уклонений, вычислить среднее по отрезку *[a, b]* значение этих квадратов и извлечь из полученного среднего квадратный корень. Тогда получится величина

$$\sqrt{\frac{1}{b-a}\int_a^b (f(x) - p_n(x))^2\, dx}, \tag{1.5}$$

которую называют *среднеквадратичным уклонением p_n от f на отрезке [a, b]*.

Заметим, что наличие в выражении (1.5) знака квадратного корня вызвано, в частности, желанием получить для среднеквадратичного уклонения ту же размерность, какую имеют значения функций. Например, если в рассматриваемой физической задаче величины *f(x)*, $p_n(x)$ измеряются в метрах, то квадрат разности этих величин имеет размерность м^2, и при отсутствии квадратного корня среднеквадратичное уклонение оказалось бы так же ве-

$$F(c_0, c_1) = \int_0^1 (x^2 + 1 - c_0 - c_1 x)^2 \, dx. \qquad (1.8)$$

Вычисления частных производных функции (1.8) приводят к следующим выражениям

$$\frac{\partial F}{\partial c_0} = \int_0^1 2(x^2 + 1 - c_0 - c_1 x)(-1)dx = -2\int_0^1 (x^2 + 1)dx + 2c_0 \int_0^1 dx + 2c_1 \int_0^1 xdx,$$

$$\frac{\partial F}{\partial c_1} = \int_0^1 2(x^2 + 1 - c_0 - c_1 x)(-x)dx = -2\int_0^1 (x^2 + 1)xdx + 2c_0 \int_0^1 xdx + 2c_1 \int_0^1 x^2 dx.$$

Если заменить здесь определенные интегралы их численными значениями и после этого приравнять полученные выражения для частных производных нулю, то для нахождения неизвестных c_0, c_1 получится система линейных алгебраических уравнений

$$c_0 + (1/2)c_1 = (4/3), \quad (1/2)c_0 + (1/3)c_1 = (3/4). \qquad (1.9)$$

Решая эту систему, получим для искомых коэффициентов значения $c_0 = 5/6$, $c_1 = 1$. Следовательно, многочленом наилучшего среднеквадратичного приближения в данном случае служит многочлен $p_1(x) = (5/6) + x$.

Предлагаем читателю выполнить самостоятельно следующее упражнение.

Упражнение 1.3. Изобразить на одном чертеже графики функции $f(x) = x^2 + 1$ и полученных в предыдущих упражнениях многочленов наилучшего среднеквадратичного приближения p_0, p_1 соответственно нулевой и первой степени для визуального решения вопроса о том, какой из этих многочленов ближе к приближаемой функции.

Замечание 1.4. Описанный выше способ приближения функций с помощью многочленов наилучшего среднеквадратичного приближения в вычислительной практике носит название *метода наименьших квадратов.*

Замечание 1.5. Аналогов метода наименьших квадратов для построения многочленов наилучшего равномерного приближения и многочленов наилучшего среднего приближения не существует. Причина этого состоит в том, что функции $F(c_0, c_1, \ldots, c_n)$, получаемые подстановкой правой части равенства (1.1) в выражения (1.3), (1.4), из-за наличия в этих выражениях операции взятия абсолютной величины *не являются дифференцируемыми по переменным* $c_0, c_1, \ldots, c_n$, то есть не имеют частных производных по указанным переменным. Этим и объясняется сложность построения многочленов наилучшего равномерного приближения и многочленов наилучшего среднего приближения.

ђqdqҷеҰ bӖ dəbӖbҮзə d єзqəоҹ ɟqdqҮdҷə

§ 2. СКАЛЯРНЫЕ ПРОИЗВЕДЕНИЯ ФУНКЦИЙ

Пусть функции v, w заданы на одном и том же отрезке $[a, b]$. Если перемножить значения этих функций в точке x, то есть рассмотреть произведение

$$z(x) = v(x)w(x), \tag{2.1}$$

то получится заданная на том же отрезке новая функция переменной x – известное из курса алгебры обычное произведение функций v, w. Если же взять среднее по отрезку $[a, b]$ значение величины (2.1), то есть среднее значение функции z по указанному отрезку, то получим число

$$\langle v, w \rangle = \frac{1}{b-a} \int_a^b v(x) w(x)\, dx, \tag{2.2}$$

которое и называется *скалярным произведением функций* v, w.

Напомним, что в естествознании *скаляром* называют величину, значениями которой являются числа; это и объясняет появление такого термина в наименовании скалярного произведения.

Отметим, что для обозначения скалярного произведения мы ввели специальный символ – острые скобки, чтобы в дальнейшем отличать скалярное произведение от обычного произведения функций.

Замечание 2.1. Скалярное произведение (2.2) можно рассматривать как обобщение понятия скалярного произведения векторов. В самом деле, функции v, w можно трактовать как бесконечномерные векторы, у которых роль индексов координат играют точки x, а роль самих координат – значения функций в указанных точках. Как и в случае конечномерных векторов, в определении (2.2) скалярного произведения функций перемножаются координаты $v(x)$, $w(x)$, отвечающие одному и тому же значению индекса x, но ввиду того, что в отличие от конечномерного случая число индексов x бесконечно, произведения координат не суммируются, а интегрируются по отрезку $[a, b]$.

Замечание 2.2. Скалярное произведение функций обладает теми же свойствами, что и скалярное произведение векторов и, более того, теми же свойствами, что и обычное произведение чисел. Это позволяет преобразовывать скалярные произведения функций по тем же правилам, по которым раскрывают скобки в школьном курсе алгебры.

Эти свойства таковы.

1) Перестановочность сомножителей:

$$\langle v, w \rangle = \langle w, v \rangle.$$

2) Возможность вынесения числового множителя за знак скалярного произведения:

$$\langle \lambda v, w \rangle = \lambda \langle v, w \rangle, \quad \langle v, \lambda w \rangle = \lambda \langle v, w \rangle.$$

3) Распределительный закон умножения относительно сложения:

$$\langle v + w, y \rangle = \langle v, y \rangle + \langle w, y \rangle, \quad \langle v, w + y \rangle = \langle v, w \rangle + \langle v, y \rangle.$$

Справедливость выписанных соотношений очевидным образом вытекает из свойств определенных интегралов.

Приступим теперь к выводу системы уравнений для нахождения коэффициентов многочлена наилучшего среднеквадратичного приближения.

Отметим, что мы не случайно ввели в рассмотрение скалярное произведение функций. В самом деле, если сопоставить формулу (1.5) с формулой (2.2), то легко убедиться в том, что подкоренное выражение F в формуле (1.5), которое минимизируется в методе наименьших квадратов, может быть представлено в виде скалярного произведения

$$F = \left\langle f - p_n, f - p_n \right\rangle. \tag{2.3}$$

Выразим величину F через коэффициенты приближающего многочлена (1.1). Для этого рассмотрим базисные функции

$$\varphi_0(x) \equiv 1,\ \varphi_1(x) = x,\ \varphi_2(x) = x^2,\ \ldots,\ \varphi_n(x) = x^n \tag{2.4}$$

и будем трактовать формулу (1.1) для приближающего многочлена p_n как представление этого многочлена в виде линейной комбинации

$$p_n = c_0\,\varphi_0 + c_1\,\varphi_1 + c_2\,\varphi_2 + \ldots + c_n\,\varphi_n \tag{2.5}$$

функций (2.4) с коэффициентами $c_0, c_1, \ldots, c_n$. Опираясь на свойства 1), 2), 3) скалярного произведения функций, преобразуем равенство (2.3) к виду

$$F = \left\langle f, f \right\rangle - 2\left\langle f, p_n \right\rangle + \left\langle p_n, p_n \right\rangle, \tag{2.6}$$

и подставим сюда вместо p_n линейную комбинацию (2.5). Тогда получим следующее выражение

$$F(c_0, c_1, \ldots, c_n) = \left\langle f, f \right\rangle - 2\left\langle f, \sum_{i=0}^{n} c_i\,\varphi_i \right\rangle + \left\langle \sum_{i=0}^{n} c_i\,\varphi_i, \sum_{j=0}^{n} c_j\varphi_j \right\rangle \tag{2.7}$$

для величины F как функции коэффициентов $c_0, c_1, \ldots, c_n$. Далее еще раз используем свойства 1), 2), 3) и в результате приходим к окончательному представлению

$$F(c_0, c_1, \ldots, c_n) = \left\langle f, f \right\rangle - 2\sum_{i=0}^{n} c_i \left\langle f, \varphi_i \right\rangle + \sum_{i,j=0}^{n} c_i\, c_j \left\langle \varphi_i, \varphi_j \right\rangle \tag{2.8}$$

для функции F.

Заметим, что при переходе от равенства (2.6) к равенству (2.7) мы намеренно заменили в третьем слагаемом справа один и тот же многочлен p_n

суммами с разными индексами суммирования. Сделано это было для того, чтобы иметь в дальнейшем возможность записать это слагаемое в виде двойной суммы с индексами i и j.

Приступим теперь к нахождению частных производных функции (2.8).

Пусть m – какое-либо целое в интервале от нуля до n.

Поскольку первое слагаемое в правой части (2.8) не зависит от c_m, производная от него по переменной c_m равна нулю.

Что же касается второго слагаемого в правой части (2.8), то переменную c_m в нем содержит лишь член суммы со значением индекса i, равным m. Дифференцируя по c_m выражение $-2c_m\langle f,\varphi_m\rangle$, получим в качестве производной от рассматриваемого слагаемого величину

$$-2\langle f,\varphi_m\rangle. \tag{2.9}$$

Наконец, в третьем слагаемом правой части равенства (2.8) – двойной сумме с индексами суммирования i, j – переменную c_m содержит, во-первых, член двойной суммы со значениями i, j, равными m, то есть член

$$c_m c_m\langle\varphi_m,\varphi_m\rangle = c_m^2\langle\varphi_m,\varphi_m\rangle$$

с производной по c_m, равной, очевидно, выражению

$$2c_m\langle\varphi_m,\varphi_m\rangle. \tag{2.10}$$

Во-вторых, переменную c_m содержат те слагаемые двойной суммы, у которых $i = m$, а индекс j принимает произвольные значения, отличные от m, и слагаемые, у которых, наоборот, $j = m$, а индекс i принимает всевозможные значения, отличные от m. Указанные слагаемые естественно сгруппировать в две подсуммы

$$\sum_{j=0,\, j\neq m}^{n} c_m c_j\langle\varphi_m,\varphi_j\rangle = c_m \sum_{j=0,\, j\neq m}^{n} c_j\langle\varphi_m,\varphi_j\rangle,\quad \sum_{i=0, i\neq m}^{n} c_i\, c_m\langle\varphi_i,\varphi_m\rangle = c_m \sum_{i=0, i\neq m}^{n} c_i\langle\varphi_i,\varphi_m\rangle,$$

производные от которых по переменной c_m имеют соответственно вид

$$\sum_{j=0,\, j\neq m}^{n} c_j\langle\varphi_m,\varphi_j\rangle,\quad \sum_{i=0, i\neq m}^{n} c_i\langle\varphi_i,\varphi_m\rangle. \tag{2.11}$$

Если разбить величину (2.10) на два слагаемых, включив первое из этих слагаемых в левую из сумм (2.11), а второе – в правую сумму, то производная от рассматриваемой двойной суммы по c_m совпадет с выражением

$$\sum_{j=0}^{n} c_j\langle\varphi_m,\varphi_j\rangle + \sum_{i=0}^{n} c_i\langle\varphi_i,\varphi_m\rangle. \tag{2.12}$$

Наконец, если в первой из фигурирующих в выражении (2.12) сумм изменить в скалярных произведениях порядок сомножителей, а затем заменить в этой сумме индекс суммирования j на индекс суммирования i, то в качестве окончательного представления для производной по c_m от фигурирующей в формуле (2.8) двойной суммы получим выражение

$$2\sum_{i=0}^{n} c_i \langle \varphi_i, \varphi_m \rangle.$$

Добавляя к этому выражению производную (2.9) от второго слагаемого в формуле (2.8) и учитывая, что производная от первого слагаемого равна нулю, приходим к равенству

$$\partial F / \partial c_m = -2\langle f, \varphi_m \rangle + 2\sum_{i=0}^{n} c_i \langle \varphi_i, \varphi_m \rangle. \tag{2.13}$$

Теорема 2.3. Коэффициенты $c_0, c_1, \ldots, c_n$ многочлена p_n наилучшего среднеквадратичного приближения для функции f на отрезке $[a, b]$ удовлетворяют системе линейных алгебраических уравнений

$$\sum_{i=0}^{n} c_i \langle \varphi_i, \varphi_m \rangle = \langle f, \varphi_m \rangle, \quad m = 0, 1, \ldots, n. \tag{2.14}$$

Доказательство. По определению многочлена наилучшего среднеквадратичного приближения набор его коэффициентов $c_0, c_1, \ldots, c_n$ есть точка минимума функции $F(c_0, c_1, \ldots, c_n)$. Как известно из курса математического анализа, необходимым условием минимума функции нескольких переменных в данной точке является обращение в этой точке в нуль всех частных производных первого порядка, которые в нашем случае задаются формулами (2.13). Приравнивая правые части этих формул нулю, сокращая полученные равенства на 2 и перенося затем скалярные произведения функции f на базисные функции φ_m в правые части равенств, приходим к системе (2.14).

Замечание 2.4. При выводе системы (2.14) мы воспользовались *необходимым* условием минимума. Более детальный анализ, проводить который в данном пособии мы не имеем возможности, показывает, что условия (2.14) являются и достаточными для того, чтобы многочлен p_n, коэффициенты $c_0, c_1, \ldots, c_n$ которого удовлетворяют этим условиям, был многочленом наилучшего среднеквадратичного приближения.

Упражнение 2.5. Найти на отрезке $[0, 1]$ многочлен первой степени наилучшего среднеквадратичного приближения для функции $f(x) = x^2 + 1$, используя систему (2.14).

Решение. В данном случае $n = 1$, поэтому система (2.14) имеет вид

$$c_0 \langle \varphi_0, \varphi_0 \rangle + c_1 \langle \varphi_1, \varphi_0 \rangle = \langle f, \varphi_0 \rangle, \quad c_0 \langle \varphi_0, \varphi_1 \rangle + c_1 \langle \varphi_1, \varphi_1 \rangle = \langle f, \varphi_1 \rangle. \tag{2.15}$$