Билет 12

> Участники:
>
> Янченков Дмитрий M3337
> Наумов Семён M3337
> Фархиев Азамат M3337
> Козырев Владислав M3337
> Комаров Андрей M3337
> Глаголев Михаил M3337

Аппроксимация функций одной переменной

1. Введение

Аппроксимация или приближенное представление одних функций другими, находит применение в алгоритмах информационно-измерительных систем. Оценка эффективности аппроксимации производится мерами различия между аппроксимируемой и аппроксимирующей функциями. Аппроксимироваться могут как детерминированные, так и случайные функции.

Целью аппроксимации является либо получение упрощенного выражения для функции, алгоритм нахождения которой весьма сложен, либо нахождение алгоритма вычисления функции, полученной или получаемой экспериментально.

2. Постановка задачи

- Дано табличное представление некоторой функции $f(x)$, которая задаёт связь между x и y на некотором множестве точек.
- Необходимо найти функцию вида $y = F(x)$, которая в точках x_i принимает значения как можно близкие к y_i, но не обязательно проходящая через эти точки $(i = 1, 2, \ldots, n)$.
- Тогда функцию $F(x)$ будем называть аппроксимирующей.

3. Практический смысл

В ходе различных экспериментов и расчётов точная зависимость между x и y может быть неизвестна, поэтому она чаще всего представлена в виде таблицы.

Но могут понадобиться и другие значения y, отличные от тех, что представлены в таблице множеством y_i. Для этого и строится такая функция $F(x)$, которая задаёт некоторое приближение.

То есть задача об аппроксимации функции состоит в том, чтобы аппроксимировать функцию, заданную таблицей.

4. Неприменимость интерполяции

Можно найти функцию методом интерполяции, тогда полученная функция будет не просто приближена к табличным значениям, а проходить через все заданные точки.

Но совпадение во всех точках функции, полученной интерполяцией, может вовсе не означать совпадение характеров исходной функции, заданной таблицей, и полученной.

К тому же измерения часто происходят с помощью различных приборов, имеющих определённую погрешность (сглаживание случайных ошибок).

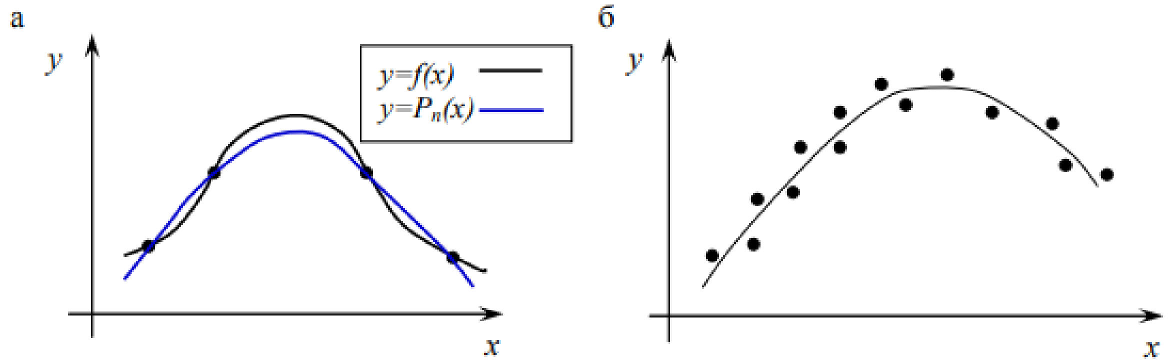
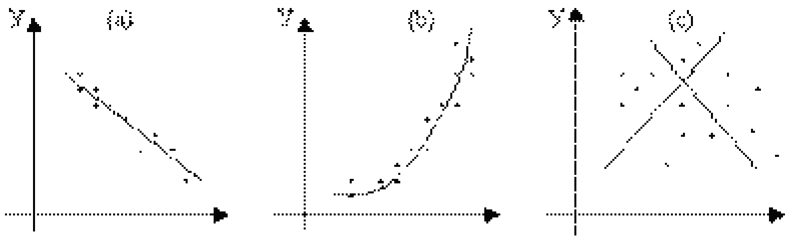


Рис. 1. Графическая интерпретация принципа построения интерполяционного полинома (а) и аппроксимирующей линии (б) для точечно заданной функции

5. Построение эмпирической формулы

1. Подбор общего вида формулы.

   Если вид функции неизвестен из физических соображений, то он угадывается геометрически. Точки, заданные в таблице, наносятся на график и угадывается общий вид функции путём сравнения с уже известными.

2. Определение значений параметров аппроксимирующей функции.

Метод наименьших квадратов

Метод наименьших квадратов (МНК) – аппроксимирующую кривую следует провести так, чтобы сумма квадратов её отклонений от табличных значений по всем узловым точкам была минимальна.

Линейный парный регрессионный анализ

Найдём линейную функцию, аппроксимирующую значения в таблице, записав их в следующем виде:

$$y = b_0 + b_1 * x \tag{1}$$

Рассчитаем сумму квадратов отклонений аппроксимирующей функции от табличных значений:

$$U = \sum_{i=1}^{n}(y_i - y(x_i))^2 = \sum_{i=1}^{n}(y_i - b_0 + b_1 * x_i)^2 \quad (2)$$

Для минимизации U необходимо приравнять нулю частные производные:

$$\frac{\partial U}{\partial b_0} = \frac{\partial U}{\partial b_1} = 0, \quad (3)$$

что эквивалентно

$$\frac{\partial U}{\partial b_0} = 2 * \sum_{i=1}^{n}(y_i - b_0 + b_1 * x_i) * (-1) \quad (4)$$
$$\frac{\partial U}{\partial b_1} = 2 * \sum i = 1^n (y_i - b_0 + b_1 * x_i) * (-x_i),$$

что эквивалентно системе линейных уравнений:

$$\begin{cases} b_0 n + b_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \\ b_0 \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} x_i^2 = \sum x_i y_i \end{cases} \quad (5)$$

Решение системы:

$$b_1 = \frac{\sum_{i=1}^{n} x_i * \sum_{i=1}^{n} y_i - n * \sum_{i=1}^{n} x_i * y_i}{(\sum_{i=1}^{n} x_i)^2 - n * \sum_{i=1}^{n} x_i^2} \quad (6)$$

$$b_0 = \frac{1}{n} * (\sum_{i=1}^{n} y_i - b_1 * \sum_{i=1}^{n} x_i) \quad (7)$$

Зная коэффициенты, можно (если нужно) вычислить и величину R (например, для сравнения различных аппроксимирующих функций). Следует помнить, что при изменении даже одного значения исходных данных (или пары значений xi, y, или одного из них) все коэффициенты изменят в общем случае свои значения, так как они полностью определяются исходными данными. Поэтому при повторении аппроксимации с несколько изменившимися данными (например, вследствие погрешностей измерения, помех, влияния неучтенных факторов и т. п.) получится другая аппроксимирующая функция, отличающаяся коэффициентами.

Пример

Пусть $n = 5$

	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$	$\sum_{i=1}^{5}$
xi	0	1	2	4	5	12
y_i	2.1	2.4	2.6	2.8	3	12.9
$x_i y_i$	0	2.4	5.2	11.2	15	33.8
x_i^2	0	1	4	16	25	46

Тогда по формулам **(6)** и **(7)**:

$$b_1 \approx 0.165 \quad (8)$$
$$b_0 \approx 2.184$$

Получаем аппроксимирующую функцию $y = 0.165x + 2.184$

Гиперболическая регрессия

$$y = b_0 + \frac{b_1}{x} \tag{9}$$

$$\begin{cases} b_0 n + b_1 \sum_{i=1}^{n} \frac{1}{x_i} = \sum_{i=1}^{n} y_i \\ b_0 \sum_{i=1}^{n} \frac{1}{x_i} + b_1 \sum_{i=1}^{n} \frac{1}{x_i^2} = \sum \frac{y_i}{x_i} \end{cases} \tag{10}$$

Степенная регрессия

$$y = b_0 * x^{b_1} \tag{11}$$

$$b_1 = \frac{\sum_{i=1}^{n} \ln x_i * \sum_{i=1}^{n} \ln y_i - n * \sum i = 1^n \ln x_i * \ln y_i}{(\sum i = 1^n \ln x_i)^2 - n * \sum i = 1^n (\ln x_i)^2} \tag{12}$$

$$b_0 = \exp(\frac{1}{n} * (\sum_{i=1}^{n} \ln y_i - b1 * \sum_{i=1}^{n} \ln x_i))$$

Показательная регрессия

$$y = a * b^x \tag{13}$$

$$\begin{cases} n * \lg a + \lg b * \sum_{i=1}^{n} x_i \sum_{i=1}^{n} \lg y_i \\ \lg a * \sum_{i=1}^{n} x_i + \lg b * \sum_{i=1}^{m} x_i^2 = \sum_{i=1}^{n} (x_i * \lg y_i) \end{cases} \tag{14}$$

Экспоненциальная регрессия

$$y = b_0 * \exp(b_1 * x) \tag{15}$$

$$b_1 = \frac{\sum_{i=1}^{n} x_i * \sum_{i=1}^{n} \ln y_i - n * \sum_{i=1}^{n} x_i * \ln yi}{(\sum_{i=1}^{n} x_i)^2 - n * \sum_{i=1}^{n} (x_i)^2} \tag{16}$$

$$b_0 = \exp(\frac{1}{n} * (\sum_{i=1}^{n} \ln y_i - b_1 * \sum i = 1^n x_i))$$

Логарифмическая регрессия

$$y = a + b * \lg x \tag{17}$$

$$\begin{cases} an + b \sum_{i=1}^{n} \lg x_i = \sum_{i=1}^{n} y_i \\ a \sum_{i=1}^{n} \lg x_i + b \sum_{i=1}^{n} (\lg x_i)^2 = \sum (y_i * \lg x_i) \end{cases} \tag{18}$$

Параболическая регрессия

$$y = b_0 + b_1 * x + b_2 * x^2 \tag{19}$$

$$\begin{cases} b_0 n + b_1 \sum_{i=1}^{n} x_i + b_2 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i \\ b_0 \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} x_i^2 + b_2 \sum_{i=1}^{n} x_i^3 = \sum x_i y_i \\ b_0 \sum_{i=1}^{n} x_i^2 + b_1 \sum_{i=1}^{n} x_i^3 + b_2 \sum_{i=1}^{n} x_i^4 = \sum x_i^2 y_i \end{cases} \tag{20}$$

Полиномиальная регрессия

$$y = a_0 + a_1 * x + a_2 * x^2 + \cdots + a_m * x^m \tag{21}$$

$$\begin{cases} c_0 a_0 + c_1 a_1 + c_2 a_2 \ldots c_m a_m = d_0 \\ c_1 a_0 + c_2 a_1 + c_3 a_2 \ldots c_{m+1} a_m = d_1 \\ \ldots c_m a_0 + c_{m+1} a_1 + c_{m+2} a_2 \ldots c_{2m} a_m = d_m \end{cases} \tag{22}$$

где $c_j = \sum_{i=1}^{n} x_i^j, j = 0, 1, \ldots, 2m, d_k = \sum_{i=1}^{n} x_i^k * y_i, k = 0, 1, \ldots, m$

Пример:

Представлены данные и необходимые для решения вычисления в виде таблицы: $n = 2$

k	x_k	x_k^2	x_k^3	x_k^4	y_k	$x_k y_k$	$x_k^2 y_k$
1	80.5	6480.5	521660.13	41993640.0625	281	22620.5	1820950.25
2	77	5929	456533	35153041.0000	272	20944	1612688.00
3	70.8	5012.64	354894.91	25126559.7696	259	18337.2	1298273.76
4	56.7	3214.89	182284.26	10335517.7121	224	12700.8	720135.36
5	39.7	1576.09	62570.773	2484059.6881	186	7384.2	293152.74
6	29.9	894.01	26730.899	799253.8801	170	5083	151981.70
$\sum$	354.6	23106.88	1604673.972	115892072.1124	1392	87069.7	5897181.81

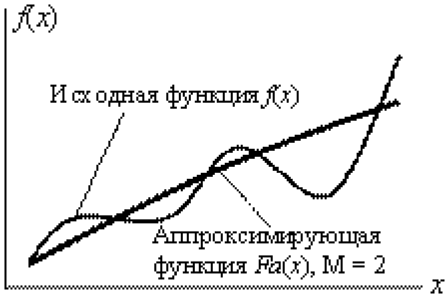
Построив и решив систему уравнений, получим:

$$\begin{gathered} a_0 = 7.425 * 10^{-3} \\ a_1 = 1.408 \\ a_2 = 120.184 \\ y = 7.425 * 10^{-3} * x^2 + 1.408 * x + 120.184 \end{gathered} \tag{23}$$

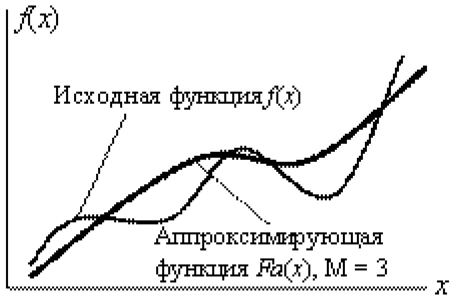
Полином степени $m < n$ обеспечивает аппроксимацию с минимальной среднеквадратичной погрешностью (обладает сглаживающими свойствами):

$E = \sqrt{\sum_{i=1}^{n} \frac{\epsilon_i^2}{n-1}}$,где $\epsilon_i = y(x_i) - \frac{1}{n} * \sum i = 1^n y_i$

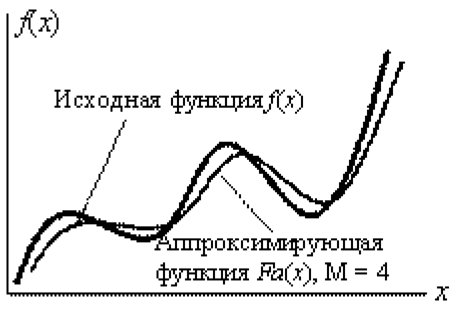
Если $m = n$ среднеквадратичная аппроксимация близка к интерполяции. Можно рассмотреть влияние степени аппроксимирующего полинома на точность аппроксимации:



а)



б)



в)