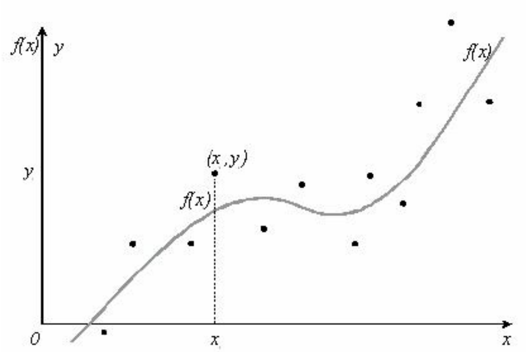
Лекция 5.
Аппроксимация функций по методу наименьших квадратов.

В инженерной деятельности часто возникает необходимость описать в виде функциональной зависимости связь между величинами, заданными таблично или в виде набора точек (x_i , y_i), где $i = 0, ..., n$. Как правило, эти табличные данные получены экспериментально и имеют погрешности.

При аппроксимации желательно получить относительно простую функциональную зависимость (например, многочлен), которая позволила бы «сгладить» экспериментальные погрешности, вычислять значения функции в точках не содержащихся в исходной таблице.

Эта функциональная зависимость должна с достаточной точностью соответствовать исходной табличной зависимости. В качестве критерия точности чаще всего используют критерий *наименьших квадратов*, т.е. определяют такую функциональную зависимость $f(x)$, при которой $R=\sum_{i=0}^{n}\left(y_i-f(x_i)\right)^2$ обращается в минимум.



Рассмотрим в качестве функциональной зависимости многочлен.

$P_m(x)=a_0+a_1x+a_2x^2+\cdots+a_mx^m$, тогда $R=\sum_{i=0}^{n}\left(y_i-P_m(x_i)\right)^2$.

Условия минимума – нулевые частные производные по всем переменным a_0, a_1, a_2, … a_m. Т.е. $\dfrac{\partial R}{\partial a_k}=-\sum_{i=0}^{n}2\left(y_i-a_0-a_1x_i-\cdots-a_mx_i\right)x_i^k=0$, или

$$\sum_{i=0}^{n}\left(y_i-a_0-a_1x_i-\cdots-a_mx_i\right)x_i^k=0,\ k=0,\ 1,\ 2,\ldots,\ m.$$

Собираем коэффициенты при неизвестных a_0, a_1, a_2, … a_m, получаем систему уравнений:

$$a_0\sum_{i=0}^{n}x_i^k+a_1\sum_{i=0}^{n}x_i^{k+1}+a_2\sum_{i=0}^{n}x_i^{k+2}+\cdots+a_m\sum_{i=0}^{n}x_i^{k+m}=\sum_{i=0}^{n}x_i^ky_i,\ k=0,\ 1,\ 2,\ldots,\ m.$$

Можно ввести обозначения:

$c_k=\sum_{i=0}^{n}x_i^k$, $b_k=\sum_{i=0}^{n}x_i^ky_i$ и переписать систему в развернутом виде.

$$\begin{cases}c_0a_0+c_1a_1+c_2a_2+\cdots c_ma_m=b_0\\ c_1a_0+c_2a_1+c_3a_2+\cdots c_{m+1}a_m=b_1\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ c_ma_0+c_{m+1}a_1+c_{m+2}a_2+\cdots c_{2m}a_m=b_m\end{cases}$$

Матрица данной системы называется матрицей Грамма. Решая эту систему линейных уравнений, получаем коэффициенты a_0, a_1, a_2, … a_m, которые являются искомыми параметрами эмпирической формулы.

Рассмотрим два частных случая $m = 1$ и $m = 2$.

1. Линейная аппроксимация (m = 1).

$P_1(x) = a_0 + a_1 x$

$$c_0 = \sum_{i=0}^{n} x_i^0 = n+1,\ c_1 = \sum_{i=0}^{n} x_i,\ c_2 = \sum_{i=0}^{n} x_i^2,\ b_0 = \sum_{i=0}^{n} y_i,\ b_1 = \sum_{i=0}^{n} x_i y_i$$

Таким образом наша система уравнений имеет вид:

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

Решаем ее методом Крамера.

$$\Delta = \begin{vmatrix} n+1 & \sum_{i=0}^{n} x_i \\ \sum_{i=0}^{n} x_i & \sum_{i=0}^{n} x_i^2 \end{vmatrix},\ a_0 = \frac{\begin{vmatrix} \sum_{i=0}^{n} y_i & \sum_{i=0}^{n} x_i \\ \sum_{i=0}^{n} x_i y_i & \sum_{i=0}^{n} x_i^2 \end{vmatrix}}{\Delta},\ a_1 = \frac{\begin{vmatrix} n+1 & \sum_{i=0}^{n} y_i \\ \sum_{i=0}^{n} x_i & \sum_{i=0}^{n} x_i y_i \end{vmatrix}}{\Delta}$$

И получаем искомую функцию $y = a_0 + a_1 x$.

2. Квадратичная аппроксимация (m = 2).

$P_2(x) = a_0 + a_1 x + a_2 x^2$

Кроме c_0, c_1, c_2, b_0, b_1 рассчитываются

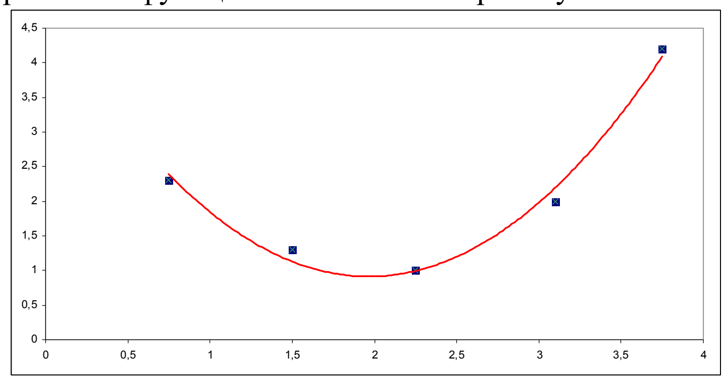
$$c_3 = \sum_{i=0}^{n} x_i^3,\ c_4 = \sum_{i=0}^{n} x_i^4,\ b_2 = \sum_{i=0}^{n} y_i x_i^2 .$$

Расширенная матрица системы уравнений: $\left(\begin{array}{ccc|c} c_0 & c_1 & c_2 & b_0 \\ c_1 & c_2 & c_3 & b_1 \\ c_2 & c_3 & c_4 & b_2 \end{array}\right)$, решив которую получим искомые коэффициенты a_0, a_1, a_2.

Пример. Получить эмпирическую формулу для функции $f(x)$, заданной таблицей, используя метод наименьших квадратов.

x	0,75	1,5	2,25	3	3,75
$f(x)$	2,3	1,3	1,0	2,2	4,2

Табличные данные изобразим на графике, из которого видно, что в качестве эмпирической функции можно взять параболу.



Найдем коэффициенты системы уравнений из таблицы.

n	x	y	x^2	x^3	x^4	xy	x^2y
0	0,75	2,3	0,5625	0,421875	0,316406	1,725	1,29375
1	1,5	1,3	2,25	3,375	5,0625	1,95	2,925
2	2,25	1	5,0625	11,39063	25,62891	2,25	5,0625
3	3	2,2	9	27	81	6,6	19,8
4	3,75	4,2	14,0625	52,73438	197,7539	15,75	59,0625
$\sum_{i=0}^{n}$	11,25	11	30,9375	94,92188	309,7617	28,275	88,14375

$$\begin{cases} 5a_0 + 11{,}25a_1 + 30{,}94a_2 = 11 \\ 11{,}25a_0 + 30{,}94a_1 + 94{,}92a_2 = 28{,}27 \\ 30{,}94a_0 + 94{,}92a_1 + 309{,}76a_2 = 88{,}14 \end{cases}$$

Решение этой системы $a_0 = 4{,}54$; $a_1 = -3{,}66$; $a_2 = 0{,}95$, и зависимость описывается формулой $y = 4{,}54 - 3{,}66x + 0{,}95x^2$.

Погрешность приближения многочленом по методу наименьших квадратов: $\Delta = \sqrt{\frac{1}{n+1}\sum_{i=0}^{n}\left(y_i - P_m(x_i)\right)^2}$

В случае нашего примера погрешность составит $\varepsilon = 0{,}01$.

Если экспериментальные точки располагаются вдоль некоторой линии, сходной по форме, например, с графиком гиперболической, показательной, логарифмической или других функций с неизвестными параметрами выбирается в качестве аппроксимирующей. Затем проводится *линеаризация* этой функции с помощью замены переменных и задача сводится к аппроксимации зависимости многочлена первой степени.

Например:

а). Показательная зависимость: $y = ab^x$, приводится к линейному виду путем логарифмирования $\ln y = \ln a + x\ln b$.

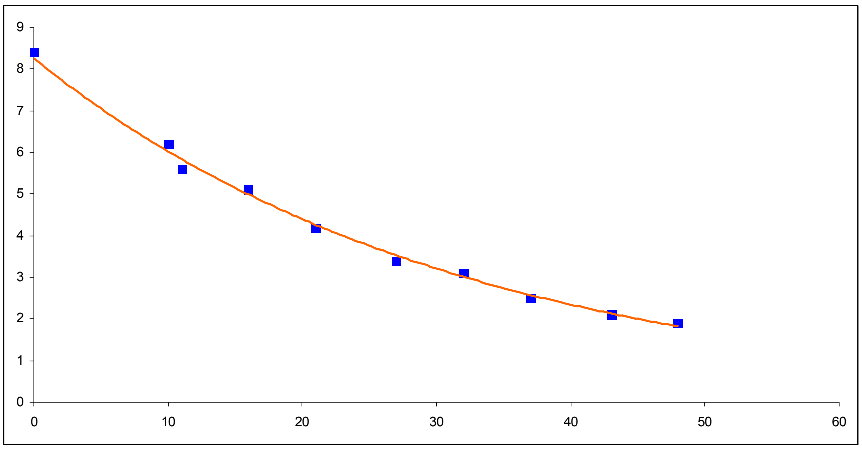
б) Степенная $y = ax^b$, аналогично $\ln y = \ln a + b\ln x$.

в) Гиперболическая $y = \frac{x}{a+bx}$, приводится к линейному виду введением новой переменной $Y = x/y$.

Пример. Результаты десяти наблюдений представлены в таблице. Установить вид зависимости между этими величинами и найти параметры эмпирической формулы.

x	0	10	11	16	21	27	32	37	43	48
y	8,4	6,2	5,6	5,1	4,2	3,4	3,1	2,5	2,1	1,9

Точечная диаграмма позволяет предположить, что зависимость между x и y экспоненциальная $y = ae^{kx}$. Вводим новые переменные $Y = lny$, $X = x$, $b = lna$, сводим зависимость к линейной $Y = kX + b$.



Перестраиваем табличную зависимость.

x	0	10	11	16	21	27	32	37	43	48
lny	2,128	1,825	1,723	1,629	1,435	1,824	1,131	0,916	0,71	0,642

По этим данным методом наименьших квадратов подберем аппроксимирующую линейную функцию. $y = a_0 + a_1x$. После решения системы получим $a_0 = 2,11$, $a_1 = -0,031 = k$. Так как $a_0 = b = lna$, то $a = e^{a0} = e^{2,11} = 8,251$.

Таким образом, полученная зависимость $y = 8,251 \cdot e^{-0,031 \cdot x}$.

Численное интегрирование.

К вычислению определенного интеграла сводятся многие практические инженерные и научные задачи. Например, вычисление площадей фигур, объемов тел, работы силы и др.

Постановка задачи: Вычислить определенный интеграл

$\int_a^b f(x)dx$, где $f(x)$ – некоторая заданная на отрезке $[a, b]$ функция.

Далеко не всегда можно вычислить интеграл в аналитическом виде по известной из курса математического анализа формуле Ньютона – Лейбница. Иногда даже при известной первообразной ее вид может оказаться очень сложным для вычислений. Также формулу Ньютона – Лейбница нельзя применить, если функция получена экспериментально в виде таблицы. Во всех этих случаях применяют численное интегрирование.

На практике наиболее широко используют квадратурные формулы:

$$\int_a^b f(x)dx \approx \sum_{i=0}^{n} A_i \cdot f(x_i)$$

где x_i – точки из отрезка $[a, b]$ (узлы квадратурной формулы), A_i – числовые коэффициенты (веса) квадратурной формулы.

За приближенное значение интеграла принимаем $\sum_{i=0}^{n} A_i \cdot f(x_i)$ - квадратурную сумму.

$$R = \int_a^b f(x)dx - \sum_{i=0}^{n} A_i \cdot f(x_i)$$ - погрешность квадратурной формулы.

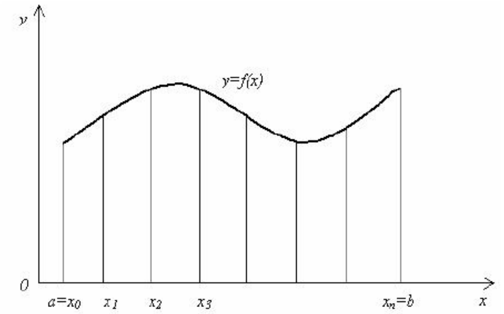
Для вывода квадратурной формулы отрезок интегрирования $[a, b]$ разбивается на n равных частичных отрезков $[x_{i-1}, x_i]$, $i = 1, \ldots, n$ длиной $h = (b - a) / h$, а интеграл $\int_a^b f(x)dx$ заменяется суммой частичных интегралов

$$\int_a^b f(x)dx \approx \sum_{i=0}^{n} \int_{x_{i-1}}^{x_i} f(x)dx .$$

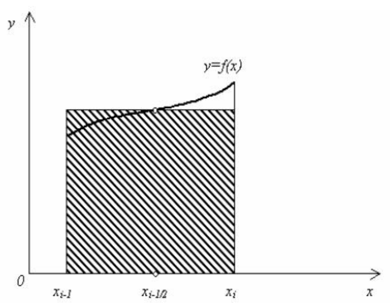
Затем подынтегральная функция $f(x)$ на частичном отрезке $[x_{i-1}, x_i]$ заменяется некоторым интерполяционным многочленом невысокой степени m $L_{m,i}(x)$, и вычисляется интеграл $\int_{x_{i-1}}^{x_i} f(x)dx = \int_{x_{i-1}}^{x_i} L_{m,i}(x)dx$ и приближенное значение интеграла будет получено в виде: $\int_a^b f(x)dx \approx \sum_{i=0}^{n} A_i \cdot f(x_i)$. В зависимости от вида интерполяционного многочлена будут получаться различные квадратурные формулы. Если будет использоваться интерполяционный многочлен нулевой степени, то будет получена формула прямоугольников, если на отрезке $[x_{i-1}, x_i]$ используется линейная интерполяция, то получается формула трапеций, если квадратичная – формула Симпсона. Но можно вывести простейшие квадратурные формулы исходя из геометрического смысла определенного интеграла, а именно находя приближенную площадь криволинейной трапеции.

$$S = \sum_{i=1}^{n} s_i = \sum_{i=1}^{n} f(x_i)\Delta x_i \quad \Delta x_i = x_i - x_{i-1}$$

Формула прямоугольников.



Можно приближенно заменить площадь криволинейной трапеции площадью ступенчатой фигуры, состоящей из n прямоугольников. Площадь i – того прямоугольника находится как произведение длины отрезка основания $[x_{i-1}, x_i]$ на значение функции в середине отрезка. $s_i \approx h \cdot f(x_{i-1/2})$



Суммируя площади всех элементарных трапеций, получим ***квадратурную формулу средних прямоугольников***.

$$\int_a^b f(x)dx \approx \sum_{i=1}^{n} f\left(x_{i-1/2}\right)h = h\left(y_{1/2} + y_{3/2} + \ldots + y_{n-3/2} + y_{n-1/2}\right)$$

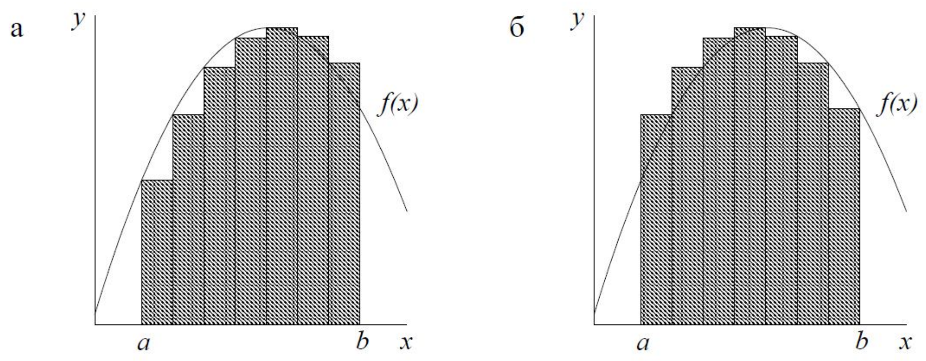
$$s_i \approx h \cdot f\left(x_i\right) \approx h \cdot f\left(x_{i-1}\right) \approx h \cdot f\left(x_{i-1/2}\right)$$

Иногда используют формулы ***правых и левых прямоугольников***:

$$\int_a^b f(x)dx \approx \sum_{i=1}^{n} f\left(x_i\right)h = h\left(y_1 + y_2 + \ldots + y_{n-1} + y_n\right)$$

$$\int_a^b f(x)dx \approx \sum_{i=1}^{n} f\left(x_{i-1}\right)h = h\left(y_0 + y_1 + \ldots + y_{n-2} + y_{n-1}\right)$$

Геометрическая иллюстрация этих формул.



Формула трапеций.

Формулу трапеций можно также получить из геометрических соображений. Отрезок интегрирования $[a, b]$ разбивается на n частей длины $h = \dfrac{b-a}{n}$. (h - шаг разбиения.)

$x_1, x_2, \cdots x_n$ - абсциссы точек деления и соответствующие ординаты кривой $y_1, y_2, \ldots, y_n$ соединяются ломаной линией. В результате построения криволинейная трапеция разбивается на ряд вертикальных полосок одной и той же ширины h, каждую из которых приближенно можно принять за трапецию. Площадь элементарной трапеции:

$$s_i = \frac{h}{2}\left(f\left(x_{i-1}\right) + f\left(x_i\right)\right)$$

Суммируя площади этих трапеций, получаем формулу:

$$\int_a^b f(x)dx \approx h\left(\frac{y_0}{2} + y_1 + y_2 + \cdots + y_{n-1} + \frac{y_n}{2}\right)$$ (Формула трапеций.)

Пример 1. Вычислить приближенно $\int_0^1 \sqrt{1+x^2}\,dx = I$.

Разобьем промежуток интегрирования на 10 частей ($n = 10$), следовательно $h = 0{,}1$.

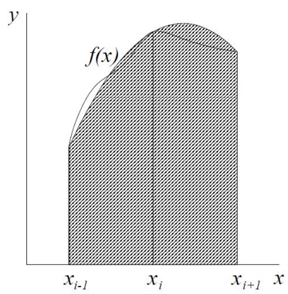
Абсциссы точек деления x_i и соответствующие им ординаты $y_i = \sqrt{1+x_i^2}$, запишем в таблице. Причем для удобства в начальной и конечной точке умножим значение на $\dfrac{1}{2}$.

i	x_i	y_i
0	0,0	0,5000*
1	0,1	1,0050
2	0,2	1,0198
3	0,3	1,0440
4	0,4	1,0770

5	0,5	1,1180
6	0,6	1,1662
7	0,7	1,2207
8	0,8	1,2806
9	0,9	1,3454
10	1,0	0,7071*

Находим $\sum_{i=0}^{10} y_i = 11{,}4838$. И по формуле трапеций имеем $I \approx 1{,}148$. Точное значение этого же интеграла, полученное по формуле Ньютона-Лейбница $I = \frac{1}{2}\sqrt{2} + \frac{1}{2}\ln(1+\sqrt{2}) \approx 1{,}1479$.

Формула Симпсона.



Более точную формулу можно получить, если профиль криволинейной полоски считать параболой (используя многочлен Лагранжа второй степени), а не прямой линией как в формуле трапеций. В этом случае можно получить формулу Симпсона:

$$\int_a^b f(x)dx \approx \frac{b-a}{3n}\{f(a) + 2[f(a+2h) + \cdots + f(a+(n-2)h)] +$$

$$+ 4[f(a+h) + \cdots + f(a+(n-1)h] + f(b)\},\ h = \frac{b-a}{n}.$$

Пример 2. Вычислить приближенно $\int_0^1 \frac{1}{1+x^2}dx = I$.

Промежуток интегрирования разбиваем на 10 частей ($n = 10$), $h = 0{,}1$.

В таблицу запишем абсциссы точек деления x_i и соответствующие им ординаты $y_i = \frac{1}{1+x_i^2}$.

i	x_i	y_i
0	0,0	1,00000
1	0,1	0,99010
2	0,2	0,96153
3	0,3	0,91743
4	0,4	0,86206
5	0,5	0,80000
6	0,6	0,73529
7	0,7	0,67114
8	0,8	0,60975
9	0,9	0,55249
10	1,0	0,50000

Отдельно суммируем значения y_i, стоящие на четных местах

$\sum(четн) = 0{,}96153 + 0{,}86206 + 0{,}73529 + 0{,}60975 = 3{,}16863$ и на нечетных местах

$\sum(нечетн) = 0{,}9901 + 0{,}9173 + 0{,}8 + 0{,}67114 + 0{,}55249 = 3{,}93116$

$$I \approx \frac{1}{30}\left[1{,}0 + 2 \cdot 3{,}16863 + 4 \cdot 3{,}93116 + 0{,}5\right] = 0{,}78537$$

Точное значение интеграла $\int_0^1 \frac{1}{1+x^2}dx = \frac{\pi}{4} \approx 0{,}7853$.

Погрешности квадратурных формул.

Для практической оценки погрешности интеграла используется правило Рунге. Для этого проводят вычисления с шагом h и $h/2$ и получают приближенные значения интегралов I_h и $I_{h/2}$.

За погрешность приближенного значения интеграла принимают величину: $(I_{h/2} - I_h)/3$ - для формулы прямоугольников и трапеций, $(I_{h/2} - I_h)/15$ - по формуле Симпсона.

В адаптивных алгоритмах автоматически определяют величину шага h, таким образом, чтобы результат удовлетворял заданной точности. За окончательные значения интеграла принимают: $I_{h/2} + (I_{h/2} - I_h)/3$ - по формуле прямоугольников, $I_{h/2} - (I_{h/2} - I_h)/3$ - формуле трапеций, $I_{h/2} - (I_{h/2} - I_h)/15$ - формуле Симпсона.