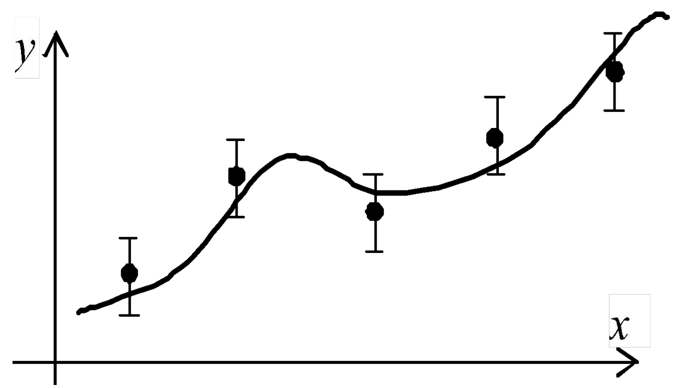
Метод наименьших квадратов (МНК)

ПОСТАНОВКА ЗАДАЧИ (МНК)

Если исходные данные в узлах интерполяции x_i, $i = 1,\ldots,N$ получены в результате опытных измерений с некоторой погрешностью ε, то точного выполнения условий интерполяции не требуется.

В этих случаях для интерполирующей функции $F(x)$ необходимо лишь приближенное выполнение условий интерполяции: $|F(x_i) - f_i| < \varepsilon$.

Данное условие означает, что интерполирующая функция $F(x)$ проходит не точно через заданные точки, а в некоторой их окрестности.



ПОСТАНОВКА ЗАДАЧИ (МНК)

Будем искать интерполирующую функцию в виде полинома, например, 3-ей степени:

$$P_3(x)=a_1+a_2x+a_3x^2+a_4x^3$$

Существует много таких полиномов, каждый из которых определяется своим набором коэффициентов (a_1, a_2, a_3, a_4).

Суть *метода наименьших квадратов* (**МНК**) состоит в том, что среди всех возможных полиномов этого вида выбирается тот, который имеет наименьшую сумму квадратов отклонений в узлах интерполяции от заданных значений.

ПОСТАНОВКА ЗАДАЧИ (МНК)

В i-й точке полином $P_3(x)$ отклоняется от значения f_i на величину $(P_3(x_i)–f_i)$. Суммируя квадраты отклонений полинома по всем точкам $i=1,…,N$, получим *функционал квадратов отклонений*:

$$G(a_1,a_2,a_3,a_4)=\sum_{i=1}^{N}(P_3(x_i)-f_i)^2=\sum_{i=1}^{N}(a_1+a_2x_i+a_3x_i^2+a_4x_i^3-f_i)^2$$

Найдем минимум этого функционала. Для этого приравняем к нулю его частные производные по переменным a_1, a_2, a_3, a_4.

ПОСТАНОВКА ЗАДАЧИ (МНК)

Используя стандартные правила дифференцирования, получим:

$$\frac{\partial G}{\partial a_1} = 2\sum_{i=1}^{N}(a_1 + a_2 x_i + a_3 x_i^2 + a_4 x_i^3 - f_i) = 0,$$

$$\frac{\partial G}{\partial a_2} = 2\sum_{i=1}^{N}x_i(a_1 + a_2 x_i + a_3 x_i^2 + a_4 x_i^3 - f_i) = 0,$$

$$\frac{\partial G}{\partial a_3} = 2\sum_{i=1}^{N}x_i^2(a_1 + a_2 x_i + a_3 x_i^2 + a_4 x_i^3 - f_i) = 0,$$

$$\frac{\partial G}{\partial a_4} = 2\sum_{i=1}^{N}x_i^3(a_1 + a_2 x_i + a_3 x_i^2 + a_4 x_i^3 - f_i) = 0.$$

ПОСТАНОВКА ЗАДАЧИ (МНК)

Собирая коэффициенты при неизвестных a_i, получим СЛАУ, которая называется ***нормальной системой***.

$$N \cdot a_1 + \sum_{i=1}^{N} x_i \cdot a_2 + \sum_{i=1}^{N} x_i^2 \cdot a_3 + \sum_{i=1}^{N} x_i^3 \cdot a_4 = \sum_{i=1}^{N} f_i,$$

$$\sum_{i=1}^{N} x_i \cdot a_1 + \sum_{i=1}^{N} x_i^2 \cdot a_2 + \sum_{i=1}^{N} x_i^3 \cdot a_3 + \sum_{i=1}^{N} x_i^4 \cdot a_4 = \sum_{i=1}^{N} f_i \cdot x_i,$$

$$\sum_{i=1}^{N} x_i^2 \cdot a_1 + \sum_{i=1}^{N} x_i^3 \cdot a_2 + \sum_{i=1}^{N} x_i^4 \cdot a_3 + \sum_{i=1}^{N} x_i^5 \cdot a_4 = \sum_{i=1}^{N} f_i \cdot x_i^2,$$

$$\sum_{i=1}^{N} x_i^3 \cdot a_1 + \sum_{i=1}^{N} x_i^4 \cdot a_2 + \sum_{i=1}^{N} x_i^5 \cdot a_3 + \sum_{i=1}^{N} x_i^6 \cdot a_4 = \sum_{i=1}^{N} f_i \cdot x_i^3.$$

Решая СЛАУ одним из известных методов, находим неизвестные коэффициенты (a_1, a_2, a_3, a_4).

МНК (пример)

Для заданной системы точек (узлов интерполяции) x_i, $i = 1,\ldots,N$ построить полином 1-ой степени, имеющий в узлах интерполяции минимальное отклонение от заданных значений f_i.

Полином 1-ой степени представляет собой линейную зависимость вида: $P_1(x)=a_1+a_2x$.

Исходные данные

$$x := \begin{pmatrix} -1 \\ -0.5 \\ 0.7 \\ 1 \\ 1.2 \end{pmatrix} \qquad f := \begin{pmatrix} -0.5 \\ -0.2 \\ 0.2 \\ 0.5 \\ 1 \end{pmatrix}$$

Решение

$$\text{ORIGIN} := 1$$

Нормальная система

$$N\cdot a_1 + \sum_{i=1}^{N} x_i\cdot a_2 = \sum_{i=1}^{N} f_i$$

$$\sum_{i=1}^{N} x_i\cdot a_1 + \sum_{i=1}^{N} \left(x_i\right)^2\cdot a_2 = \sum_{i=1}^{N} \left(x_i\cdot f_i\right)$$

МНК (пример)

Вычислим коэффициенты при неизвестных a_1, a_2 и свободные члены:

$$N := 5$$

$$\sum_{i=1}^{N} x_i = 1.4 \qquad \sum_{i=1}^{N} \left(x_i\right)^2 = 4.18$$

$$\sum_{i=1}^{N} f_i = 1 \qquad \sum_{i=1}^{N} \left(x_i \cdot f_i\right) = 2.44$$

Решим нормальную систему методом обратной матрицы:

$$A := \begin{pmatrix} 5 & 1.4 \\ 1.4 & 4.18 \end{pmatrix} \qquad b := \begin{pmatrix} 1 \\ 2.44 \end{pmatrix} \qquad a := A^{-1} \cdot b = \begin{pmatrix} 0.04 \\ 0.57 \end{pmatrix}$$

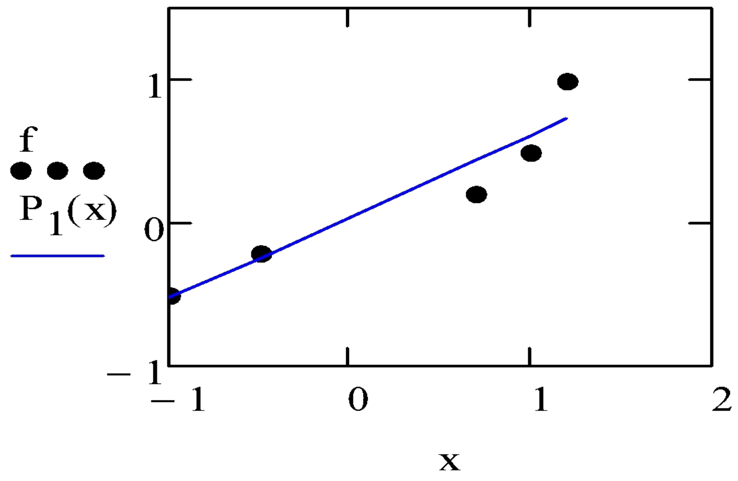
МНК (пример)

Вид полинома: $P_1(x) := 0.04 + 0.57 \cdot x$

Сумма квадратов отклонений:

$$\sum_{i=1}^{N} \left(P_1\left(x_i\right) - f_i\right)^2 = 0.148$$

График функции:



Численное интегрирование

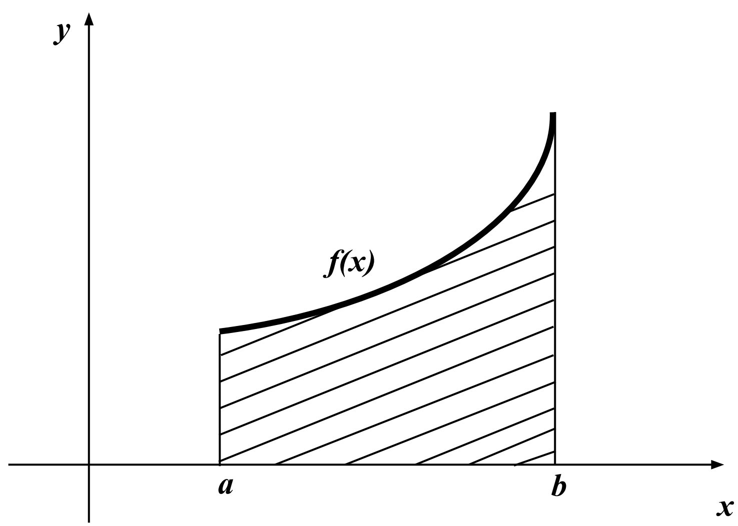
ПОСТАНОВКА ЗАДАЧИ

Найти значение определенного интеграла

$$I = \int_a^b f(x)dx$$

для функции $f(x)$, заданной на некотором отрезке $[a, b]$.

Исходя из геометрической интерпретации определенного интеграла, основой методов численного интегрирования является нахождение площади криволинейной трапеции, ограниченной подынтегральной функцией $f(x)$, осью x, прямыми $x=a$ и $x=b$.



ПОСТАНОВКА ЗАДАЧИ

Формулы приближенного интегрирования называются ***квадратурными формулами***. Простейшей квадратурной формулой является ***общая формула прямоугольников***, которая вычисляется с помощью приближенного равенства:

$$I_n \cong \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1})$$

где x_i, $i = 0, 1, \ldots, n$ – заданная система точек на отрезке интегрирования

$[a, b]$; ξ_i – произвольная точка элементарного промежутка $[x_{i-1}, x_i]$.

Геометрически это означает, что площадь криволинейной трапеции заменяется на сумму площадей прямоугольников, основанием которых является *i*-й интервал $[x_{i-1}, x_i]$, а высотой – значение функции $f(\xi_i)$.

Погрешность любой квадратурной формулы определяется модулем разности между значением, вычисленным по квадратурной формуле I_n, и точным значением интеграла I:

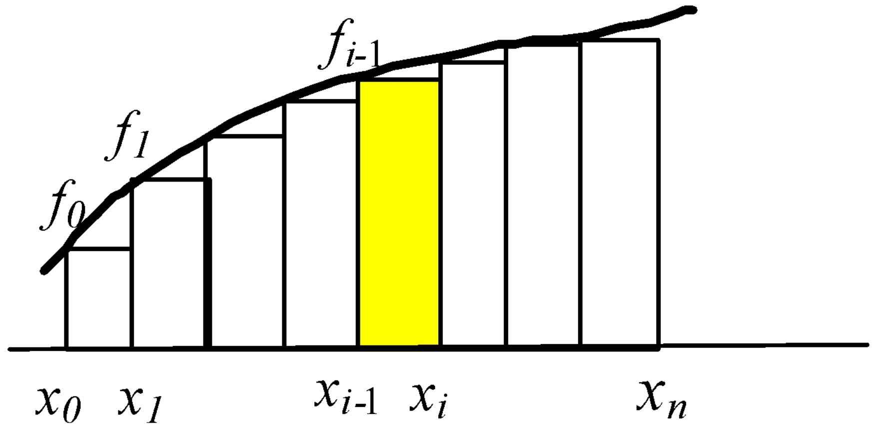
$$|\Psi_n| = |I_n - I|$$

Квадратурные формулы

- Формула левых прямоугольников
- Формула правых прямоугольников
- Формула средних прямоугольников
- Формула трапеций
- Формула Симпсона

Формула левых прямоугольников

Аппроксимируем подынтегральную функцию левой кусочно-постоянной интерполяцией, т.е. $f(\xi_i)=f(x_{i-1})$.



$$I_{l\_pr}=f_0(x_1-x_0)+f_1(x_2-x_1)+..+f_{n-1}(x_n-x_{n-1})=\sum_{i=1}^{n} f_{i-1}(x_i-x_{i-1})$$

При равномерной сетке длина интервала: $h=x_i-x_{i-1}=\dfrac{b-a}{n}$

Получим формулу: $I_{l\_pr}=h\sum_{i=1}^{n} f(x_{i-1})=h\sum_{i=0}^{n-1} f(x_i)$

Формула левых прямоугольников (пример)

Дан определенный интеграл: $\int_1^2 (\frac{4}{x} - 5x^4 + 2\sqrt{x})dx$

Вычислить значение интеграла в пакете MathCad с помощью *оператора интегрирования.*

Используя формулу левых прямоугольников при сетке с количеством отрезков *n*=10 составить **П-Ф** и вычислить приближенное значение интеграла. Оценить погрешность.

$ORIGIN := 0$

$f(x) := \frac{4}{x} - 5 \cdot x^4 + 2 \cdot \sqrt{x}$

$$\boxed{I_0 := \int_1^2 f(x)\,dx = -25.79}$$

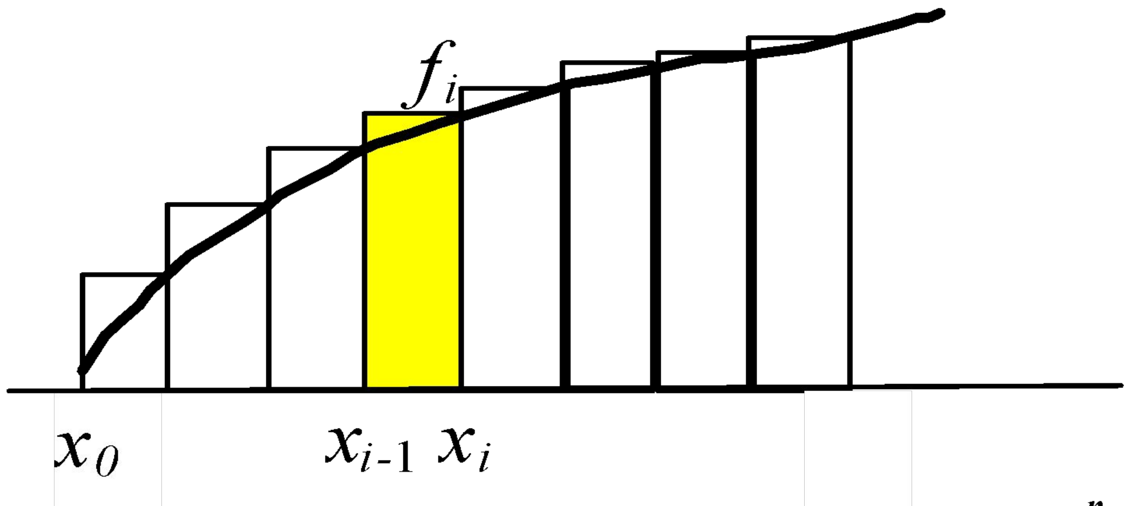
$$I(f,a,b,n) := \left| \begin{array}{l} h \leftarrow \frac{(b-a)}{n} \\ S \leftarrow 0 \\ \text{for } i \in 0..n-1 \\ \quad \left| \begin{array}{l} x_i \leftarrow a + i \cdot h \\ S \leftarrow S + f(x_i) \end{array} \right. \\ S \leftarrow h \cdot S \\ S \end{array} \right.$$

$I_{1\_pr} := I(f,1,2,10) = -22.095$

$\psi := \left| I_{1\_pr} - I_0 \right| = 3.694$

Формула правых прямоугольников

Аппроксимируем подынтегральную функцию правой кусочно-постоянной интерполяцией, т.е. $f(\xi_i) = f(x_i)$.



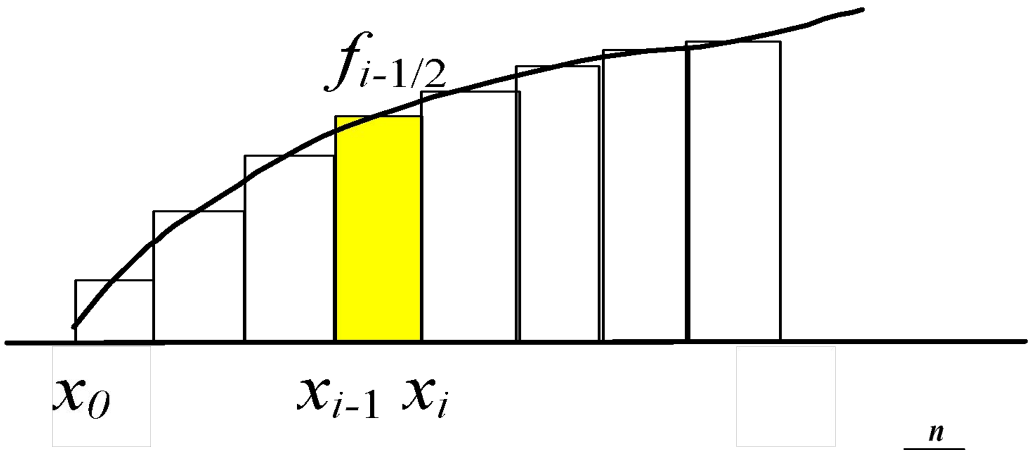
$$I_{p\_pr} = f_1(x_1 - x_0) + f_2(x_2 - x_1) + .. + f_n(x_n - x_{n-1}) = \sum_{i=1}^{n} f_i(x_i - x_{i-1})$$

При равномерной сетке получим формулу:

$$I_{p\_pr} = h\sum_{i=1}^{n} f(x_i)$$

Формула средних прямоугольников

Заменим на каждом локальном отрезке $[x_{i-1}, x_i]$ значение подынтегральной функции на ее значение в середине интервала, т.е. $f(\xi_i) = f(\frac{x_{i-1}+x_i}{2}) = f_{i-1/2}$



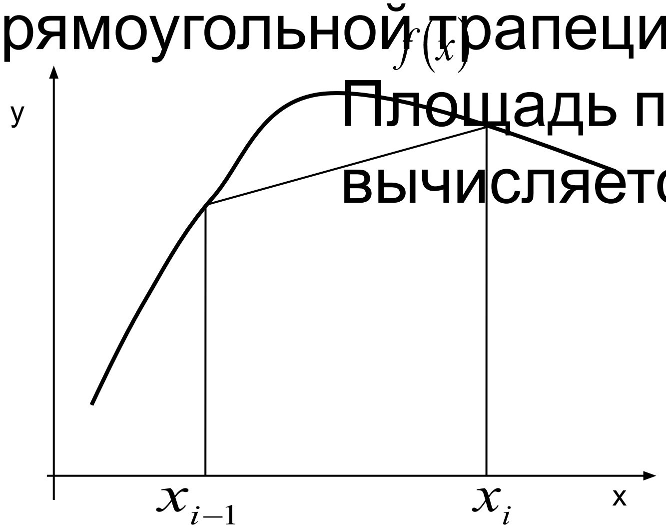
$$I_{c\_pr} = f_{1/2}(x_1 - x_0) + f_{3/2}(x_2 - x_1) + .. + f_{n-1/2}(x_n - x_{n-1}) = \sum_{i=1}^{n} f_{i-1/2}(x_i - x_{i-1})$$

При равномерной сетке : $f_{i-1/2} = f(x_i - \frac{h}{2})$

Получим формулу: $I_{c\_pr} = h\sum_{i=1}^{n} f(x_i - \frac{h}{2}) = h\sum_{i=0}^{n-1} f(x_i + \frac{h}{2})$

Формула трапеций

На каждом локальном отрезке [x_{i-1}, x_i] аппроксимируем подынтегральную функцию линейной зависимостью (кусочно-линейная интерполяция). В этом случае криволинейная трапеция заменяется прямоугольной трапецией.



Площадь прямоугольной трапеции вычисляется по формуле:

$$\int_{x_{i-1}}^{x_i} f(x)dx \cong \frac{f(x_{i-1}) + f(x_i)}{2}(x_i - x_{i-1})$$

Суммируя площади всех трапеций при равномерной сетке получим формулу:

$$I_{tr} = h\left[\frac{f(x_0) + f(x_n)}{2} + \sum_{i=1}^{n-1} f(x_i)\right]$$

Формула Симпсона

На каждом локальном отрезке $[x_{i-1}, x_i]$ аппроксимируем подынтегральную функцию кусочно-параболической зависимостью, т.е. параболой, проходящей через три точки:

$$\left(x_{i-1}, f\left(x_{i-1}\right)\right), \left(x_{i-1/2}, f\left(x_{i-1/2}\right)\right), \left(x_i, f\left(x_i\right)\right)$$

где $x_{i-1/2} = \dfrac{x_{i-1} + x_i}{2}$ – середина отрезка.

Интерполируя на каждом отрезке подынтегральную функцию полиномом Лагранжа второй степени при равномерной сетке получим формулу:

$$I_{simp} = \frac{h}{6} \sum_{i=1}^{n} \left[f\left(x_{i-1}\right) + 4f\left(\frac{x_{i-1} + x_i}{2}\right) + f\left(x_i\right) \right]$$