

Całkowanie metodą warstwową

Metody Monte Carlo w Fizyce

Julia Potempa (411073)



Akademia Górniczo-Hutnicza
im. Stanisława Staszica w Krakowie

9. kwietnia 2025r.

Cel ćwiczenia

Porównanie metod całkowania Monte Carlo warstwowych (optymalną i nieoptymalną) z prostym całkowaniem.

1 Wstęp

Proste całkowanie

Całka postaci

$$C = \int_a^b g(x) dx$$

ma funkcję gęstości prawdopodobieństwa $f(x) = \text{const}$, która wykorzystując warunek normalizacji może zostać zapisana jako:

$$f(x) = \frac{1}{b-a}$$

zatem całkę C można zapisać jako:

$$C = \int_a^b \frac{g(x)}{f(x)} f(x) dx = \int_a^b [(b-a)g(x)] f(x)$$

i jej wartość można przybliżyć średnią z próby

$$C \approx \bar{g} = \frac{1}{N} \sum_{i=1}^N (b-a)g(x), \quad x_i \sim U(a, b)$$

Losowanie z rozkładu jednorodnego w zakresie $[a, b]$ można zapisać jako $x_i = a + (b-a)U_i$, $U_i \sim U(0, 1)$. Następnie można policzyć drugi moment oraz wariancję średniej:

$$\overline{g^2} = \frac{1}{N} \sum_{i=1}^N [(b-a)g(x)]^2$$

$$\sigma_{\bar{g}}^2 = \frac{\overline{g^2} - \bar{g}^2}{N}$$

Metoda losowania systematycznego (losowanie warstwowe nieoptymalne)

Metoda losowania systematycznego polega na podzieleniu przedziału całkowania na M podprzedziałów, w których ilość losowań jest skalowana prawdopodobieństwem wylosowania zmiennej w danym podprzedziale. Przy ustalonej identycznej szerokości każdego przedziału ($\Delta x = (b-a)/M$) granice m -tego przedziału przybierają postać:

$$x_m = a + \Delta x \cdot (m-1)$$

$$x_{m+1} = x_m + \Delta x$$

Prawdopodobieństwo ogólnie można zapisać jako

$$p_m = \int_{x_m}^{x_{m+1}} f(x) dx$$

natomiast dla równomiernego podziału i jednorodnego rozkładu w tym przedziale $p_m = 1/M$. Liczba losowań zatem przyjmuje postać:

$$N_m = p_m \cdot N$$

Pierwszy i drugi ($n = 1, 2$) moment oraz wariancję wyrażamy odpowiednio jako:

$$\bar{g}_m^n = \frac{1}{N_m} \sum_{i=1}^{N_m} [(b-a)g(x_{i_m})]^n, \quad x_{i_m} \sim U(x_m, x_{m+1})$$

$$\sigma_m^2 = \bar{g}_m^2 - (\bar{g}_m)^2$$

Ostateczna wartość całki można oszacować jako średnią i dodatkowo można obliczyć wariancję średniej:

$$C \approx \bar{g} = \sum_{m=1}^M p_m \cdot \bar{g}_m$$

$$\sigma_{\bar{g}^2} = \sum_{m=1}^M \frac{p_m^2}{Nm} \cdot \sigma_m^2$$

Metoda losowania warstwowego (losowanie optymalne)

Metoda losowania warstwowego przebiega analogicznie do metody systematycznej. Zmienia się tylko sposób dobierania ilości losowań w każdym z przedziałów.

$$N_m = \frac{p_m \hat{\sigma}_m}{\sum_{j=1}^M p_j \hat{\sigma}_j} \cdot N$$

gdzie $\hat{\sigma}_j$ to wartości odchylenia standardowego, które obliczamy metodą nieoptymalną dla małych wartości N (100, 1000). W trakcie wykonywania kolejnych iteracji wyznaczamy "dokładniejsze" wartości σ_m i na ich podstawie liczymy wariancję średniej.

2 Metodyka

Wszystkie trzy zastosowane metody zaimplementowano w języku C++, a wykresy zostały wygenerowane za pomocą kodu w języku Python. Zastosowano następujące parametry:

- całka $C_1 = \int_{-3}^3 (1 + \tanh(x)) dx = 6$,

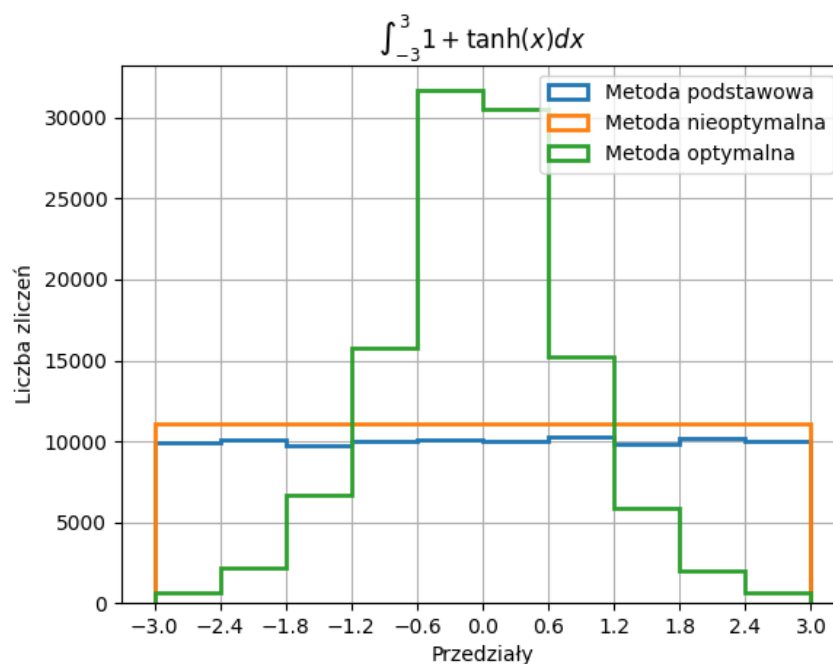
- całka $C_2 = \int_0^{10} \frac{1}{1+x^2} dx = \arctan(10) - \arctan(0)$,
- całka $C_3 = \int_0^1 \cos^{10}(\pi x) dx = 0,24609375$,
- punkty zapisu $N = 10^k$, $k = 2, 3, 4, 5$,
- ilość podprzedziałów $M = 10$.

Dla metody warstwowej przy liczbie losowań 100 estymowano wariancję metodą nieoptymalną dla 100 losowań, natomiast przy 1000 losowań lub więcej, estymowano metodą nieoptymalną dla 1000 losowań. Dodatkowo dla każdej z metod został wyliczony błąd względny dany wzorem:

$$R = \frac{\sigma_{\bar{g}}}{\bar{g}} \cdot 100\%$$

3 Wyniki

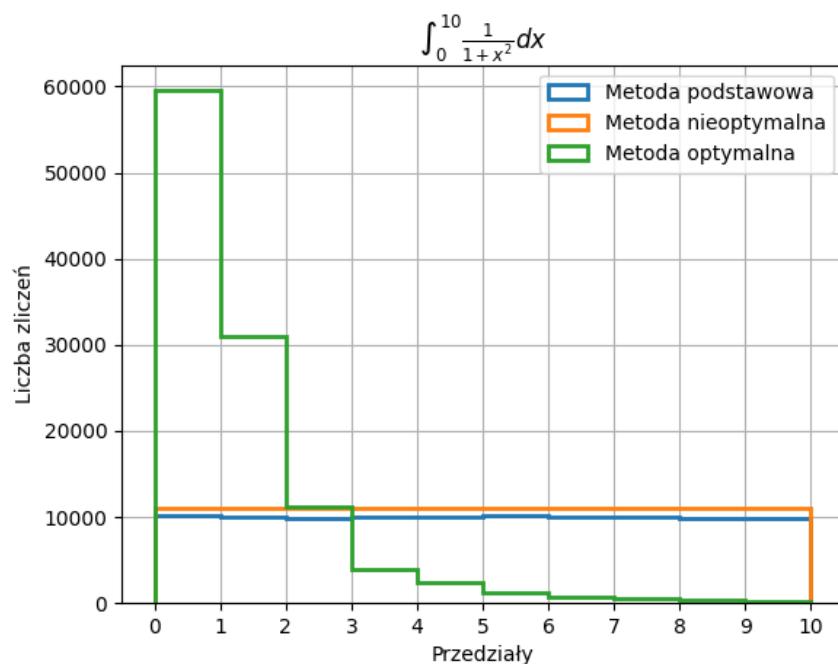
Uzyskane wyniki zostały zwizualizowane za pomocą zarówno wykresów prezentujących histogramy jak i tabel z liczbą zliczeń w danych numerach przedziałów dla metody optymalnej.



Rysunek 1: Histogram wartości dla całki z funkcji $f(x) = 1 + \tanh x$.

Tabela 1: Histogram wartości dla całki z funkcji $f(x) = 1 + \tanh x$.

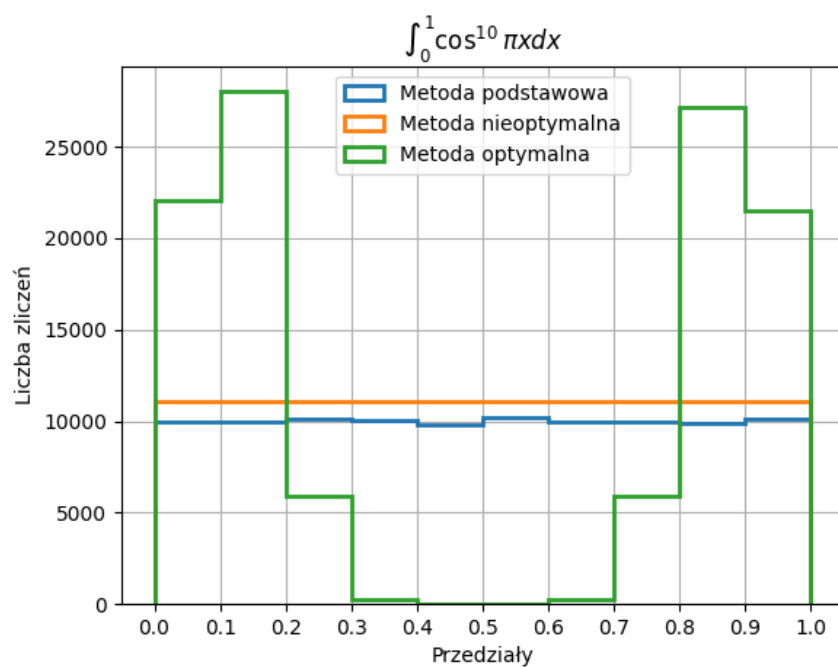
Nr przedziału	1	2	3	4	5	6	7	8	9	10
Zliczenia	610	2143	6647	15754	31676	30519	15263	5874	1993	640



Rysunek 2: Histogram wartości dla całki z funkcji $f(x) = \frac{1}{1+x^2}$.

Tabela 2: Histogram wartości dla całki z funkcji $f(x) = \frac{1}{1+x^2}$.

Nr przedziału	1	2	3	4	5	6	7	8	9	10
Zliczenia	59539	30953	11280	3940	2331	1265	715	516	329	254

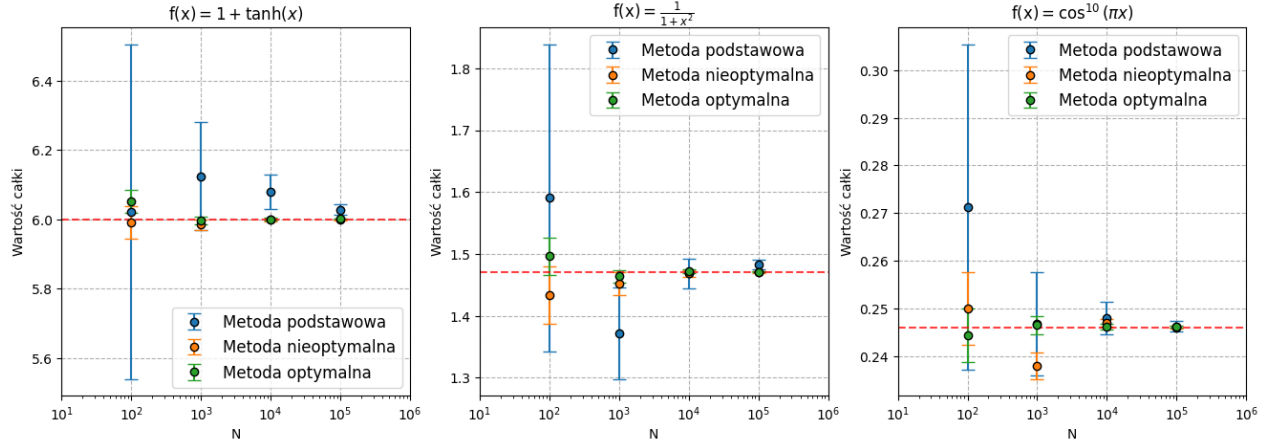


Rysunek 3: Histogram wartości dla całki z funkcji $f(x) = \cos^{10}(\pi x)$.

Tabela 3: Histogram wartości dla całki z funkcji $f(x) = \cos^{10}(\pi x)$.

Nr przedziału	1	2	3	4	5	6	7	8	9	10
Zliczenia	22096	28018	5941	227	4	4	241	5898	27193	21504

Wyniki statystyczne zostały podsumowane na wykresie wraz ze słupkami odchylenia standardowego. Dodatkowo przykładowe wartości całki, odchylenie standardowe oraz błąd względny zostały zebrane dla funkcji $f(x) = 1 + \tanh x$ w tabelach, osobno dla każdej metody.



Rysunek 4: Zmiana wartości obliczanej całki w funkcji liczby powtórzeń N wraz z zaznaczoną linią wskazującą poprawny wynik całkowania.

Tabela 4: Wartości statystyczne dla metody podstawowej liczenia całki z $f(x) = 1 + \tanh x$.

Liczba powtórzeń N	Wartość całki	Odchylenie standardowe	Błąd względny [%]
100	6,02277	0,48265	8,01375
1000	6,12457	0,155192	2,53392
10000	6,07985	0,0490531	0,806814
100000	6,02775	0,0154984	0,257118

Tabela 5: Wartości statystyczne dla metody nieoptymalnej (systematycznej) liczenia całki z $f(x) = 1 + \tanh x$.

Liczba powtórzeń N	Wartość całki	Odchylenie standardowe	Błąd względny [%]
100	5,99256	0,0467185	0,779608
1000	5,98481	0,0158483	0,264809
10000	5,99888	0,00491794	0,0819809
100000	5,99875	0,00153693	0,0256208

Tabela 6: Wartości statystyczne dla metody optymalnej (warstwowej) liczenia całki z $f(x) = 1 + \tanh x$.

Liczba powtórzeń N	Wartość całki	Odchylenie standardowe	Błąd względny [%]
100	6,05132	0,033073	0,546541
1000	5,99672	0,01093	0,182267
10000	6,00095	0,00344089	0,057339
100000	6,00215	0,00109478	0,0182397

4 Wnioski

Podsumowując, w niniejszym sprawozdaniu zostały porównane trzy metody całkowania: podstawową, systematyczną i warstwową. Najbardziej optymalną metodą okazała się metoda warstwowa, w której dzięki estymowanej wartości wariancji, najszybciej dało się naprowadzić program na odpowiednią wartość całki. Niestety dodatkowo wymaga to zaimplementowania metody nieoptymalnej. Rozkłady wartości zostały zaprezentowane na histogramach, które pokazały dużą ilość losowanych wartości w przedziałach dużej zmienności całkowanej funkcji.