Monte Carlo: całkowanie metodą warstwową

11 kwietnia 2022

1 Wstęp

W projekcie należy użyć metody MC do oszacowania wartości całek

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

$$C_2 = \int_{0}^{10} \frac{1}{1+x^2} dx = arctan(10) - arctan(0) \tag{2}$$

$$C_3 = \int_{0}^{1} \cos^{10}(\pi x)dx = 0.24609375 \tag{3}$$

Całki te należy oszacować używając: 1) metody podstawowej, 2) losowania systematycznego, 3) losowania warstwowego.

1.1 Metoda podstawowa

Dla całki postaci

$$C = \int_a^b g(x)dx \tag{4}$$

identyfikujemy fgp jako $f(x) = const$, z warunku normalizacji dostajemy

$$\int_a^b f(x)dx = const \int_a^b 1\, dx = const(b-a) = 1 \qquad \rightarrow \qquad f(x) = \frac{1}{b-a} \tag{5}$$

Modyfikujemy całkę

$$C = \int_a^b \frac{g(x)}{f(x)} f(x)dx = \int_a^b [(b-a)g(x)]f(x) \tag{6}$$

a jej wartość przybliżamy średnią z próby

$$C \approx \bar{g} = \frac{1}{N}\sum_{i=1}^{N}(b-a)\cdot g(x_i), \qquad x_i \sim U(a,b) \tag{7}$$

gdzie losowanie z rozkładu jednorodnego w zakresie $[a, b]$, wykonujemy stosując prostą transformację $x_i = a + (b-a)\cdot U_i, \quad U_i \sim U(0,1)$. $U(0,1)$ to generator liczb pseudolosowych o rozkładzie jednorodnym w przedziale $(0, 1)$. Liczymy jeszcze drugi moment

$$\overline{g^2} = \frac{1}{N}\sum_{i=1}^{N}\left[(b-a)\cdot g(x_i)\right]^2, \qquad x_i \sim U(a,b) \tag{8}$$

i wariancję średniej

$$\sigma^2_{\bar{g}} = \frac{\overline{g^2} - \bar{g}^2}{N} \tag{9}$$

1.2 metoda losowania systematycznego (warstwowe nieoptymalne)

Najpierw dokonujemy podziału obszaru całkowania na M podobszarów. Załóżmy, że mają identyczną szerokość $\Delta x = (b-a)/M$. Wówczas lewą (x_m) i prawą (x_{m+1}) granicę przedziału wyznaczają

$$x_m = a + \Delta x \cdot (m-1), \qquad m = 1, 2, \dots, M \tag{10}$$

$$x_{m+1} = x_m + \Delta x \tag{11}$$

W metodzie losowania systematycznego (warstwowego nieoptymalnego) określamy prawdopodobieństwo wylosowania zmiennej z danego podprzedziału p_m

$$p_m = \int_a^b f(x)dx \tag{12}$$

co dla **równomiernego podziału i jednorodnego fozkładu fgp** daje $p_m = 1/M$. Dla każdego podprzedziału m-tego określamy liczbę losowań

$$N_m = p_m \cdot N \tag{13}$$

obliczamy $n = 1$ i 2 moment oraz wariancję

$$\overline{g^n}_m = \frac{1}{N_m}\sum_{i=1}^{N_m}[(b-a)\cdot g(x_{i_m})]^n, \qquad x_{i_m} \sim U(x_m, x_{m+1}) \tag{14}$$

$$\sigma_m^2 = \overline{g^2}_m - (\bar{g}_m)^2 \tag{15}$$

Teraz możemy oszacować wartość całki C jako średnią i wariancję średniej

$$C \approx \bar{g} = \sum_{m=1}^{M} p_m \cdot \bar{g}_m \tag{16}$$

$$\sigma_{\bar{g}}^2 = \sum_{m=1}^{M} \frac{p_m^2}{N_m} \cdot \sigma_m^2 \tag{17}$$

1.3 metoda losowania warstwowego (optymalnego)

W metodzie tej postępujemy identycznie jak dla losowania systematycznego poza jednym wyjątkiem, liczbę losowań N_m w każdym podprzedziale określamy według wzoru

$$N_m = \frac{p_m \widehat{\sigma}_m}{\sum_{j=1}^{M} p_j \widehat{\sigma}_j} N \tag{18}$$

gdzie: $\widehat{\sigma}_j$ to prognozowane/szacowane wartości odchylenia standardowego, które obliczamy metodą podstawową dla małej wartości N (np. $N = 10^2, 10^3$) - bo dokładnych wartości nie znamy. Oczywiście w trakcie wykonywania właściwych obliczeń (metoda warstwowa) na bieżąco wyznaczamy "dokładniejsze" wartości σ_m i ich ostatecznie używamy do liczenia wariancji średniej.

2 Zadania do wykonania

1. Zaimplementować metodę podstawową całkowania MC i oszacować wartość całki C_1, odchylenie standardowe średniej $\sigma_{\bar{g}}$ oraz błąd względny $R = (\sigma_{\bar{g}}/\bar{g}) \cdot 100\%$ dla $N = 10^k, \quad k = 2, 3, 4, 5$. Przedział $[a, b]$ podzielić na $M = 10$ podprzedziałów o identycznej szerokości. Dla każdego N określić ilość liczb wpadających do każdego podprzedziału. Wyniki zebrać w tabelce.

2. Zaimplementować metodę losowania systematycznego i powtórzyć obliczenia z zadania 1.

3. Zaimplementować metodę losowania warstwowego i powtórzyć obliczenia z zadania 1. Szacunkowe wartości $\widehat{\sigma}_m$ określić przy użyciu metody podstawowej, używając 100 losowań (metoda podstawowa) dla $N = 10^2$ (metoda warstwowa) oraz 1000 losowań (metoda podstawowa) dla $N \geqslant 10^3$ (metoda warstwowa). (Należy pamiętać, że do szacowania $\widehat{\sigma}_m$ stosujemy krótkie serie losowań.)

4. Powtórzyć obliczenia z punktów: 1, 2 i 3 dla całek C_2 i C_3

5. Tabelki z wynikami zamieścić w raporcie, wyniki przeanalizować i skomentować, zamieścić przykładowe (wybrane) histogramy rozkładu ilości losowań dla poszczególnych podprzedziałów, krótko skomentować (uzasadnić) różnice w wysokościach słupków na histogramach dla trzech użytych metod.