Analiza matematyczna dla informatyków.

Mieczysław Cichoń, ver. 3.2/2023

Mieczysław Cichoń - WMI UAM

Przypuśćmy, że mamy znaleźć całkę $\int_a^b f(x)dx$, gdzie $f(x)$ jest pewną funkcją ciągłą określoną w przedziale $[a, b]$. Poprzednio podaliśmy wiele przykładów obliczania takich całek za pomocą funkcji pierwotnych.

Zauważmy jednak, że wszystkie metody daje się zastosować jedynie do dość wąskiej klasy całek. Poza tą klasą musimy się uciekać do metod rachunku przybliżonego (całki Riemanna).

Obecnie poznamy najprostsze z metod, w których wzory przybliżone na całkę wykorzystują pewną wartość funkcji podcałkowej, obliczane dla pewnych wartości zmiennej niezależnej. Wzory te otrzymuje się najczęściej z równań geometrycznych, traktując całkę oznaczoną $\int_a^b f(x)dx$ jako pole pewnej figury geometrycznej ograniczonej krzywą $f(x)$.

Metoda trapezów.

Jeżeli funkcja f jest ciągła na przedziale domkniętym $[a, b]$ oraz jeżeli π jest podziałem przedziału $[a, b]$ takim, że $\pi: \; a = x_0 < x_1 < \ldots < x_n = b$, to

$$\int_a^b f(x)dx \approx \frac{b-a}{2n}[f(x_0)+2f(x_1)+2f(x_2)+\ldots+2f(x_{n-1})+f(x_n)]. \, (*)$$

Funkcja f jest funkcją ciągłą na przedziale $[a, b]$. Dokonajmy podziału przedziału $[a, b] \; : \; a = x_0 < x_1 < \ldots < x_n = b$. Przez P_k oznaczamy punkt, który powstał z przecięcia krzywej $y = f(x)$ oraz prostej $x = x_k, \;\; k = 0, 2, \ldots, n$. Przez Δx oznaczamy $\Delta x = x_i - x_{i-1}$. Łączymy łamaną kolejne punkty $P_0, P_1, \ldots, P_n$.

Zróbmy rysunek: powstają trapezy. Korzystając ze znanego z geometrii wzoru na pole trapezu $P = \frac{1}{2}(a+b) \cdot h$, gdzie a, b są podstawami trapezu, a h jego wysokością obliczmy pole $\mid P^i \mid$ trapezu P^i

$$\mid P^i \mid = \frac{\Delta x}{2}\left[f(x_{i-1} + f(x_i)\right] .$$

Pole obszaru jest sumą pól wszystkich trapezów, a więc

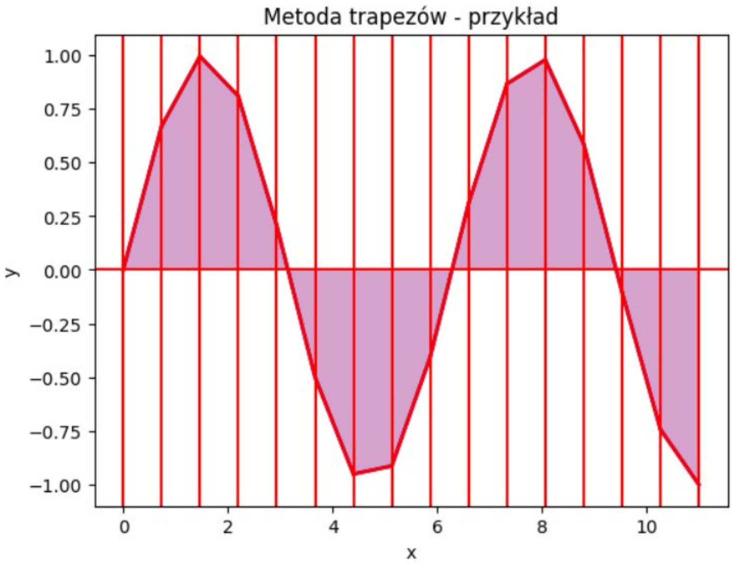
$$\mid P \mid = \mid P^1 \mid + \ldots + \mid P^n \mid = \frac{\Delta x}{2}(f(x_0) + f(x_1)) + \frac{\Delta x}{2}(f(x_1) + f(x_2)) + \ldots +$$
$$+ \frac{\Delta x}{2}(f(x_{n-1}) + 2f(x_n)) =$$
$$= \frac{\Delta x}{2}\left[f(x_0) + 2(x_1) + 2f(x_2) + \ldots + 2f(x_{n-1}) + f(x_n)\right] .$$

Łatwo zauważyć, że otrzymaliśmy taką samą sumę jak we wzorze $(*)$.

Tak więc całka $\int_a^b f(x)dx$ tzn. obszar ograniczony krzywą $y = f(x)$, osią Ox oraz rzędnymi $x = a$, $x = b$ została obliczona z pewnym przybliżeniem. Oczywiście wynik ten obarczony jest błędem. Poniżej podamy jaki błąd maksymalny może wystąpić w przypadku obliczania całki oznaczonej $\int_a^b f(x)dx$ przy pomocy metody trapezów.

Błąd przybliżenia dla metody trapezów. $(**)$

Jeżeli M jest dodatnią liczbą rzeczywistą taką, że $| f''(x) | < M$ dla każdego $x \in [a, b]$, to błąd jaki pojawia się w przypadku obliczania całki oznaczonej $\int_a^b f(x)dx$ metodą trapezów jest nie większy niż $\dfrac{M(b-a)^3}{12n^2}$.



Metoda trapezów - przykład
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
y
0
2
4
6
8
10
x

Przykład 12.

Wykorzystując metodę trapezów, przybliżyć całkę $\int_1^2 \frac{1}{x} dx$, dla $n = 10$. Podać błąd przybliżenia.

Ponieważ $\frac{b-a}{2n} = (2-1) \cdot \frac{1}{20} = 0,05$, więc wykorzystując wzór $(*)$ otrzymujemy

$$\int_1^2 \frac{1}{x} dx \approx (0,05) \cdot (13,8754) \approx 0.6938 \ .$$

Obliczamy błąd jaki wystąpił w powyższych obliczeniach.
Obliczamy drugą pochodną funkcji $f(x) = \frac{1}{x}$.

$$f'(x) = -\frac{1}{x^2} \ , \qquad\qquad f''(x) = \frac{2}{x^3} \ .$$

Aby obliczyć błąd maksymalny musimy tak dobrać stałą M, aby $| f''(x) | < M$. Ponieważ maksymalną wartością $f''(x) = \frac{2}{x^3}$ w przedziale $[1, 2]$ jest $f''(1) = \frac{2}{1^3} = 2$, więc możemy przyjąć $M = 2$. Stąd na mocy $(**)$ maksymalny błąd obliczenia jaki wystąpił przy stosowniu metody trapezów jest nie większy niż wartość

$$\frac{M(b-a)^3}{12n^2} = \frac{2 \cdot (2-1)^3}{12 \cdot (10)^2} = \frac{1}{600} < 2 \cdot 10^{-3} \ .$$

Metoda Simpsona.

Jeżeli f jest funkcją ciągłą określoną w przedziale $[a, b]$ oraz jeżeli π jest podziałem przedziału $[a, b]$ takim, że
$\pi:\ a = x_0 < x_1 < \ldots < x_n = b$, to

$$\int_a^b f(x)dx \approx \frac{b-a}{3n}\Big[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \ldots + \\ + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)\Big]\ .$$

Rozważmy funkcję postaci $y = cx^2 + dx + e$. Jeżeli $c \neq 0$, to wykresem tej funkcji jest parabola.

Pole obszaru **pod parabolą** wynosi

$$A = \int_{-h}^{h}(cx^2 + dx + e)dx = \left[c\frac{x^3}{3} + d\frac{x^2}{2} + ex\right]_{-h}^{h} = \frac{1}{3}h(2ch^2 + 6e)\ .$$

Dla jasności: skąd otrzymamy parabolę? Przypomnijmy: przez 3 punkty nie będące współliniowe przechodzi dokładnie jedna parabola. Będziemy więc brać **3 kolejne** punkty podziału np. P_0, P_1, P_2 i znajdziemy wzór paraboli

$$y - cx^2 + dx + e$$

przez nie przechodzący.

A jeśli są współliniowe? Ten przypadek też jest objęty dla $c = d = 0$. Ale zauważmy: procedura działa poprawnie, gdy liczba punktów podziału jest postaci od 0 do $2N$. Przypomnijmy, że przedziały mają u nas stałą długość h.

Ponieważ punkty P_0, P_1, P_2 leżą na paraboli więc współrzędne tych punktów $P_0(-h, y_0)$, $P_1(0, y_1)$, $P_2(h, y_2)$ spełniają równanie tej paraboli tzn. $y_0 = ch^2 - dh + e$, $y_1 = e$, $y_2 = ch^2 + dh + e$.

Stąd $y_0 + 4y_1 + y_2 = 2ch^2 + 6e$. W konsekwencji

$$A = \frac{h}{3}(y_0 + 4y_1 + y_2).$$

Pole całego obszaru ograniczonego prostymi $x = x_0$ i $x = x_2$ wynosi więc

$$B = \int_{x_0}^{x_2} (cx^2 + dx + e)dx = \left[c\frac{x^3}{3} + d\frac{x^2}{2} + ex \right]_{x_0}^{x_2} =$$

$$= c\frac{x_2^3}{3} + d\frac{x_2^2}{2} + ex_2 - c\frac{x_0^3}{3} - d\frac{x_0^2}{2} - ex_0 =$$

$$= \frac{c}{3}(x_2^3 - x_0^3) + \frac{d}{2}(x_2^2 - x_0^2) + e(x_2 - x_0) =$$

$$= \frac{c}{3}(x_2 - x_0)(x_2^2 - x_2x_0 + x_0^2) + \frac{d}{2}(x_2 - x_0)(x_2 + x_0) + e(x_2 - x_0) =$$

$$= (x_2 - x_0)\left[\frac{c}{3}(x_2^2 - x_2x_0 + x_0^2) + \frac{d}{2}(x_2 + x_0) + e \right].$$

Jeżeli $f(x) \geq 0$ na $[a, b]$ to w metodzie Simpsona całkę oznaczoną przybliżamy wartością pola obszaru znajdującego się pod krzywą f w przedziale $[a, b]$.

Niech $h = \dfrac{b-a}{n}$, $\pi:\ a = x_0 < x_1 < \ldots < x_n = b_n$.

Niech $P_k(x_k, y_k)$ będzie punktem na krzywej $y = f(x)$ powstałym z przecięcia wykresu funkcji f oraz prostej $x = x_k$. Jeżeli punkty P_0, P_1, P_2 leżą na paraboli $y = cx^2 + dx + e$ to pole obszaru poniżej łuku $P_0P_1P_2$ wyraża się wzorem $\frac{h}{3}(y_0 + 4y_1 + y_2)$.

Jeżeli P_2, P_3, P_4 są punktami leżącymi na paraboli $y = cx^2 + dx + e$, która przybliża wykres naszej funkcji, to pole obszaru poniżej wyraża się wzorem $\frac{h}{3}(y_2 + 4y_3 + y_4)$.

Kontynuujemy nasze postępowanie i dla punktów P_{n-2}, P_{n-1}, P_n otrzymujemy, że pole obszaru położonego poniżej paraboli przybliżającej wykres naszej funkcji od x_{n-2} do x_n wyraża się wzorem

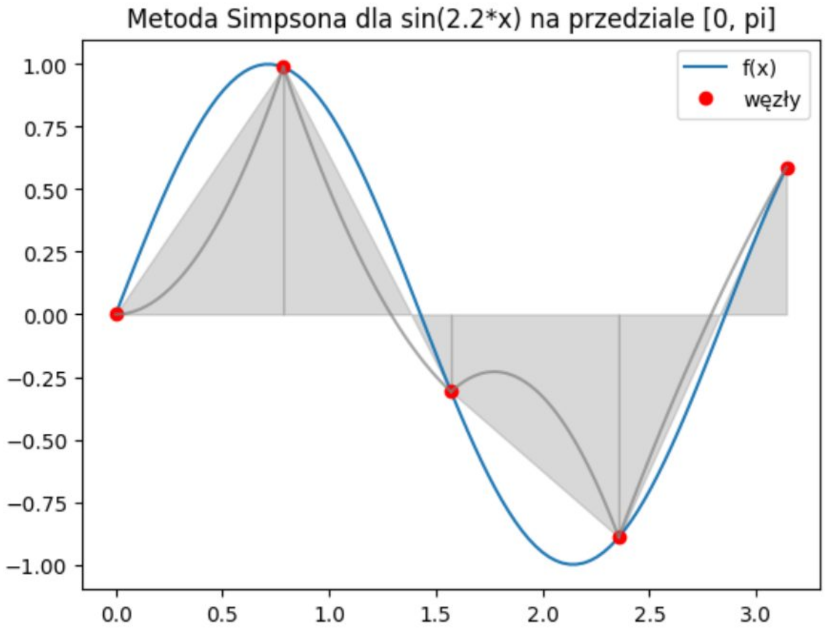
$$\frac{h}{3}(y_{n-2} + 4y_{n-1} + y_n).$$

Sumując otrzymane wyniki mamy

$$\int_a^b f(x)dx \approx \frac{h}{3}(y_0 + 4y_1 + 2y_2 + 4y_3 + \ldots + 2y_{n-2} + 4y_{n-1} + y_n) \ .$$

Błąd przybliżenia dla metody Simpsona.

Jeżeli M jest dodatnią liczbą rzeczywistą taką, że $| f^{(4)}(x) | < M$ dla każdego $x \in [a, b]$ (gdzie $f^{(4)}(x)$ oznacza czwartą pochodną funkcji f w punkcie x, to błąd jaki pojawia się w przypadku obliczania całki nieoznaczonej $\int_a^b f(x)dx$ metodą Simpsona jest nie większy niż: $\dfrac{M(b-a)^5}{180n^4}$.


Metoda Simpsona dla sin(2.2*x) na przedziale [0, pi]
f(x)
węzły
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0.0
0.5
1.0
1.5
2.0
2.5
3.0

Przykład 13.

Wykorzystując metodę Simpsona, przybliżyć całkę $\int_1^2 \frac{1}{x} dx$, dla $n = 10$.

Podać błąd przybliżenia. Ponieważ

$$\frac{b-a}{3n} = \frac{2-1}{30} = \frac{1}{30}$$

więc wykorzystując metodę Simpsona mamy

$$\int_1^2 \frac{1}{x} dx \approx \frac{1}{30} \cdot (20.7944) \approx 0.6932 \ .$$

Chcąc obliczyć błąd jaki wystąpił przy obliczaniu całki tą metodą należy znaleźć czwartą pochodną funkcji f

$$f(x) = \frac{1}{x}\ ,\ f'(x) = -\frac{1}{x^2}\ ,\ f''(x) = \frac{2}{x^3}\ ,\ f'''(x) = -\frac{6}{x^4}\ ,\ f^{(4)}(x) = \frac{24}{x^5}\ .$$

Ponieważ największą wartością pochodnej czwartego rzędu w przedziale $[1, 2]$ jest

$$f^{(4)}(1) = 24$$

więc możemy przyjąć $M = 4$.

Na mocy naszych rozważań błąd jakiego szukamy jest nie większy niż $\frac{24(2-1)^5}{180(10)^4} = \frac{2}{150000} < 1,4 \cdot 10^{-5}$.

Zauważmy, że błąd jaki wystąpił przy obliczaniu całki metodą Simpsona jest mniejszy niż w przypadku stosowania metody trapezów.

Uwaga.

Ale my przecież potrafimy obliczyć całkę! mamy

$$\int_1^2 \frac{1}{x}\, dx = \left[\ln x\right]_1^2 = \ln 2 - \ln 1 = \ln 2.$$

Obliczane numerycznie całki przybliżają więc wartość dokładną liczby niewymiernej $\ln 2$. Mamy kolejną metodę przybliżania liczb niewymiernych! Wyrażamy je jako wyniki całkowania i przybliżamy numerycznie...

Błędy metody dla kilku popularnych metod numerycznych obliczania całek:

Metoda	**Błąd**
Metoda prostokątów lewych	$O(h)$
Metoda prostokątów prawych	$O(h)$
Metoda trapezów	$O(h^2)$
Metoda Simpsona	$O(h^4)$
Metoda Gaussa gdzie n jest stopniem dokładności	$O(h^{2n+1})$,

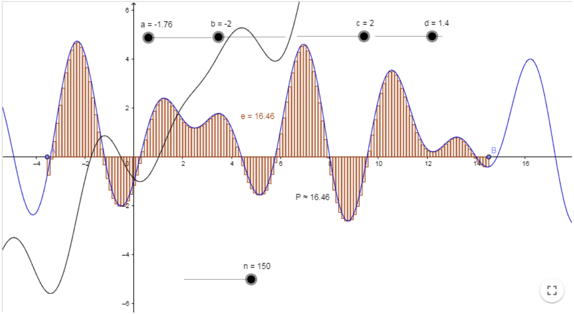
Przedstawiono tu błędy dla jednowymiarowych całek, gdzie h oznacza odległość między węzłami, a n jest stopniem dokładności odpowiednio dla metody Gaussa (o tej metodzie mozna powiedzieć dopiero po omówieniu aproksymacji funkcji).

Należy jednak zauważyć, że błędy te są szacunkowe i zależą od funkcji, której całkę obliczamy oraz od przedziału całkowania. W praktyce ważne jest również uwzględnienie dokładności reprezentacji liczb w komputerze, która wpływa na precyzję wyniku.

Geogebra.

Wracamy do wykorzystania programu "Geogebra" do ilustracji materiału. Oczywiście: proszę uruchamiać skrypty na swoich komputerach...

Tutaj sumy Riemanna i całki. Całka a pole. Kolejne materiały między innymi tutaj oraz jeszcze tu. Inne materiały: opracować lub poszukać.



Klasyczny przykład zastosowania całek w **grafice komputerowej** to równanie renderowania Kajiya (co gorsza - potrzebne na ogół metody całkowania numerycznego :-)).
Inny przypadek: w **teorii kolejkowania** - np. równanie całkowe Pollaczka (ale też: o czym później) - całki podwójne niewłaściwe (prawdopodobieństwo opóźniania kolejkowania) - to w kolejnym wykładzie (niestety tylko o ile zdążymy!).

Absolutna "klasyka" - szacowanie sum (częste w obliczeniach) poprzez całki:

$$\int_{m-1}^{n} f(x)\,dx \leq \sum_{k=m}^{n} f(k) \leq \int_{m}^{n+1} f(x)\,dx.$$

A teraz zadanie: jakie są **założenia**, aby powyższy wzór był prawdziwy? Odpowiedzi szukaj na wykładach z analizy...