Imię i Nazwisko	Przedmiot	Data oddania	Ocena
Paulina Roszkowska Aleksandra Szum	Algorytmy i struktury danych	24 stycznia 2020	
Sprawozdanie III			
Temat sprawozdania	Klątwa wielowymiarowości		

1 Opis ćwiczenia

Ćwiczenie miało na celu zapoznanie z nieintuicyjnym zachowaniem wielowymiarowej przestrzeni. Do ćwiczenia użyto algorytm Monte Carlo.

Pierwszą częścią ćwiczenia było dokonanie obliczeń teoretycznych oraz z wykorzystaniem metody Monte Carlo ilorazu objętości hiperkuli do hipersześcianu na niej opisanej dla 3, 9, 15 oraz 21 wymiaru.

Drugą częścią ćwiczenia było wyznaczenie, jaki procent masy całej pomarańczy zajmuje skórka dla 3, 9, 15 i 21 wymiaru.

2 Wstęp teoretyczny

Przekleństwo wielowymiarowości odnosi się do wielu właściwości przestrzeni wielowymiarowych. Przede wszystkim dotyczy wykładniczego wzrostu niezbędnych danych eksperymentalnych w zależności od wymiaru przestrzeni przy rozwiązywaniu problemów probabilistyczno-statystycznego rozpoznawania wzorców, uczenia maszynowego. Dotyczy to również wykładniczego wzrostu liczby wariantów w kombinatorycznych problemach w zależności od wielkości początkowych danych, co prowadzi do odpowiedniego zwiększenia złożoności algorytmów ponownego wyboru.

Funkcja odległości w matematyce to funkcja odległości lub metryka jest uogólnieniem pojęcia odległości fizycznej. Metryka jest funkcją, która zachowuje się zgodnie z określonym zestawem reguł i jest sposobem na opisanie oznacza to, że elementy przestrzeni są „blisko” lub „daleko od siebie”.

W geometrii wyższych wymiarów hipersfera to zbiór punktów w stałej odległości od danego punktu zwanego jego środkiem. Wraz ze wzrostem promienia hipersfery zmniejsza się jego krzywizna.

Hipersześcian to uogólnienie sześcianu w wymiarowych przestrzeniach kartezjańskich. Hipersześcianami są obok sześcianu między innymi odcinek i kwadrat, jednak nazwy hipersześcian używa się najczęściej dla przestrzeni o wymiarach powyżej trzech.

3 Obliczenia numeryczne

Kod znajduje się na stronie `https://github.com/aleksandraszum/AiSD_projekty/tree/master/report3`.

4 Analiza wyników

4.1 Intuicja brzegów przestrzeni

Obliczenia teoretyczne oraz uzyskane metodą Monte Carlo znajdują się w tabeli 1.

Tabela 1: Zbiorcze przedstawienie wyników uzyskanych wyznaczeniem ilorazu objętości hiperkuli i objętości hipersześcianu wyznaczając ze wzoru oraz wyznaczając z wykorzystaniem algorytmu Monte Carlo.

Wymiar n	3	9	15	21
Liczba kroków	10^3	10^3	10^6	10^7
Objętość hiperkuli	$\frac{4}{3}\pi r^3$	$\frac{32\pi^4}{945} \cdot r^9$	$\frac{256\pi^7}{2027025} \cdot r^{15}$	$\frac{2048\pi^{10}}{13749310575} \cdot r^{21}$
Objętość hipersześcianu	a^3	a^9	a^{15}	a^{21}
Wartość teoretyczna $\frac{V_{kuli}}{V_{walca}}$	$5,25 \cdot 10^{-1}$	$6,44 \cdot 10^{-3}$	$1,16 \cdot 10^{-5}$	$6,65 \cdot 10^{-9}$
Wartość wyznaczona przez algorytm	$5,36 \cdot 10^{-1}$	$6,00 \cdot 10^{-3}$	$1.20 \cdot 10^{-5}$	$0,00$
Błąd	$1,1 \cdot 10^{-3}$	$4,4 \cdot 10^{-4}$	$1,16 \cdot 10^{-5}$	$6,65 \cdot 10^{-9}$

Objętość dla n-wymiarowych hiperkul zapisuje się wzorami:

$$V_{2m} = \frac{\pi^m}{m!} \cdot r^{2m}, \tag{1}$$

gdy n jest liczbą parzystą oraz

$$V_{2m+1} = \frac{2^{2m+1} \cdot m! \cdot \pi^m}{(2m+1)!} \cdot r^{2m+1}, \tag{2}$$

gdy n jest liczbą nieparzystą.

Promień analizowanej kuli wynosił $r = 1$. Podnoszenie liczby 1 do dowolnej potęgi n nie sprawi, że uzyskany wynik będzie różny od wartości 1. Biorąc to pod uwagę, objętość analizowanej kuli nie będzie zależała od promienia. Natomiast biorąc pod uwagę wzory (1-2) zaobserwowano, że wraz ze zwiększeniem liczby wymiarów, wartość wyrażenia przez które mnożona jest odpowiednia potęga promienia, dąży do zera.

Długość boku analizowanego sześcianu wynosi $a = 2$. Objętość hipersześcianu rośnie 2-razy wraz ze wzrostem wymiaru o 1.

Wynika z tego, że wraz ze zwiększeniem się liczby wymiarów, iloraz objętości hiperkuli do objętości hipersześcianu będzie się zmniejszać, dążąc do wartości 0.

Uzyskane błędy przybliżenia są niewielkie - dla wymiaru trzeciego są trzeciego rzędu.

Zaobserwowano, że algorytm dla wymiaru rzędu 21, wyznaczył, że wartość objętości hiperkuli wynosi 0. Biorąc pod uwagę, że wartości teoretyczne można zaokrąglić do 0, jest to poprawny wynik.

4.2 Problem skórki pomarańczy

Wzór na gęstość d zapisuje się jako:

$$d = \frac{m}{V}, \tag{3}$$

gdzie m jest masą, a V objętością. Zakładając, że skórka pomarańczy ma taką samą gęstość jak jej jadalny miąższ, wyznaczenie procentu masy skórki w całej pomarańczy można dokonać wyznaczając procent objętości miąższu do całej pomarańczy.

Wartość teoretyczną wyznaczono za pomocą wzoru:

$$\texttt{Wartość\_teoretyczna} = \frac{V_{\texttt{miąższ}}}{V_{\texttt{pomarańcza}}}, \tag{4}$$

gdzie $V_{\texttt{miąższ}}$ jest objętością n-wymiarowej kuli o promieniu $0.9 \cdot r = 0,9$, a $V_{\texttt{pomarańcza}}$ jest objętością n-wymiarowej kuli o promieniu $r = 1$.

Obliczenia teoretyczne oraz uzyskane metodą Monte Carlo znajdują się w tabeli 2.

Tabela 2: Zbiorcze przedstawienie wyników uzyskanych wyznaczeniem procentowej zawartości skórki pomarańczy w pomarańczy wyznaczone metodą analityczną oraz wykorzystując algorytm Monte Carlo

Wymiar n	L. kroków	Wartość teoretyczna [%]	Wartość wyznaczona przez algorytm [%]
3	10^3	72.9	71.4
9	10^6	38.8	38.8
15	10^7	20.6	20.3
21	10^8	10.9	Error: float division by zero

Metodą Monte Carlo uzyskano wyniki tylko dla wymiaru trzeciego, dziewiątego oraz piętnastego. Zauważono, że wraz ze zwiększeniem liczby wymiarów należy zwiększać liczbę wywołań algorytmu. Ze względu na długi czas wykonywania się algorytmu, przeprowadzono analizę dla liczby wywołań kodu wynoszącej maksymalnie 10^8.

5 Wnioski

Podczas przeprowadzenia ćwiczenia wykazano, że wraz ze zwiększeniem się wymiaru, czynnik przez który mnoży się promień kuli o zadanym promieniu dąży do wartości zero.

Stosunek objętości hiperkuli do objętości hipersześcianu na niej opisanej zmniejsza się wraz ze zwiększeniem się wymiaru.

Procentowa zawartość masy skórki do masy pomarańczy zmniejsza się wraz ze zwiększeniem liczby wymiarów.

Aby uzyskać możliwie jak najdokładniejsze wartości wyznaczone przez algorytm Monte-Carlo należy zwiększać liczbę wywołań wraz ze zwiększeniem się liczby wymiarów. Dzięki tej operacji nie doprowadza się do sytuacji, kiedy następuje dzielenie przez zero. Kosztem uzyskania zbliżonego wyniku bądź braku błędu (dzielenie przez zero) jest zwiększenie czasu działania algorytmu.

6 Bibliografia

1. Wikipedia, *Hiperkula*, dostęp online:
   https://pl.wikipedia.org/wiki/Hiperkula

2. Wikipedia, *Przekleństwo wielowymiarowości* , dostęp online: https://pl.wikipedia.org/wiki/Przekle%C5%84stwo_wymiarowo%C5%9Bci

3. Wikipedia, *Curse of dimensionality* , dostęp online: https://en.wikipedia.org/wiki/Curse_of_dimensionality/