Метод Монте-Карло в паралельних обчисленнях

Пішта Я.В.
Національний технічний університет України «Київський політехнічний інститут»,
pishta.yaroslav@bk.ru

The article examined the principles of the Monte-Carlo method, and analysing the relevance of further studies of this method with the exploitation for the distributed computing

Введення

Метод статистичних випробувань - метод обчислювальної та прикладної математики, заснований на моделюванні випадкових величин і побудові статистичних оцінок для шуканих величин; те саме, що Монте-Карло методи. Прийнято вважати, що метод статистичних випробувань виник в 1944 році, коли у зв'язку з роботами по створенню атомних реакторів американські вчені Дж. фон Нейман і С. Улам почали широко застосовувати апарат теорії ймовірностей для вирішення прикладних задач за допомогою ЕОМ[1]. Спочатку метод використовувався головним чином для вирішення складних завдань теорії переносу випромінювання та нейтронної фізики, де традиційні чисельні методи виявилися мало придатними. Потім його вплив поширився на більший клас задач статистичної фізики, дуже різних за своїм змістом. Метод застосовується для розв'язання задач теорії ігор, теорії масового обслуговування та математичної економіки, завдань теорії передачі повідомлень при наявності перешкод і т.д.

Постановка завдання

Для вирішення завдання по методам Монте-Карло насамперед будують імовірнісну модель, представляють шукану величину, наприклад багатовимірний інтеграл, у вигляді математичного сподівання функціоналу від випадкового процесу, який потім моделюється на комп'ютері. У результаті проведення обчислювального експерименту отримують потрібну вибірку і результати всіх випробувань усереднюють.

Принципова математична основа використання методів Монте-Карло - Посилений закон великих чисел у формі А.Н. Колмогорова[2].

Завдання оцінювання числа π, метод Hit-Or-Miss

Метод полягає в наступному: розглянемо одиничний квадрат на координатній площині і чверть окружності одиничного радіуса з центром в початку координат (рис. 1.2).

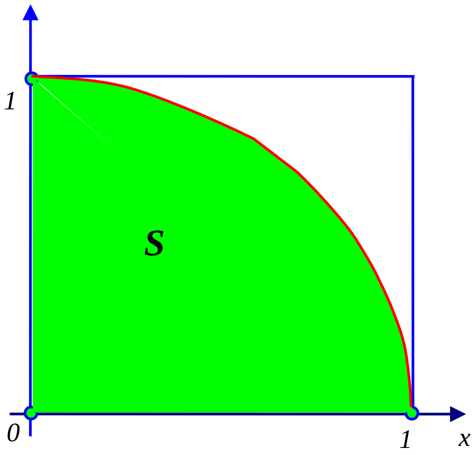

Рис. 1.2 Метод Hit-Or-Miss

Нехай одиничний експеримент полягає в тому, що в одиничному квадраті випадково вибирається будь-яка точка. Розглянемо подію A, яке у тому, що точка потрапила у розглянутий сектор кола[6].

Площа квадрата Smax = 1, площа виділеної області $S = \frac{\pi}{4}$.

Розглянемо випадкову величину $\xi_i = \begin{Bmatrix} 1, & A; \\ 0, & \overline{A} \end{Bmatrix}$ (1.1).

Введемо $S_N = \sum_{i=1}^{N} \xi_i, \quad \frac{S_N}{N} -$ (1.2) частота настання події A.

Тоді $\frac{S_N}{N} \in \left\{ \frac{0}{N}; \frac{1}{N}; \ldots; \frac{N}{N} \right\}$ (1.3), випадкова величина $\frac{S_N}{N}$ (1.4) підпорядковується біноміальному розподілу. Визначимо параметри розподілу.

$$p = P(\xi_i = 1) = \frac{\frac{\pi}{4}}{1} = \frac{\pi}{4}$$

$$q = P(\xi_i = 0) = 1 - p = 1 - \frac{\pi}{4} = \frac{3\pi}{4}$$

$$P\left(\frac{S_N}{N} = \frac{m}{N} \right) = C_N^m p^m q^{n-m}. \qquad (1.4)$$

Будемо розглядати $\frac{S_N}{N}$ – частоту настання події A - в якості оцінки величини p - імовірності того, що точка потрапила в зазначений сектор кола. Знаючи теоретичну ймовірність цієї події $\frac{\pi}{4}$, ми можемо оцінити цю ймовірність по досить великій вибірці ($\frac{S_N}{N}$)та розглянути наступне рівняння: $\frac{S_N}{N} = \frac{\pi}{4}$ (1.5)

Тоді можна розглянути $\hat{\pi}_N = \frac{4S_N}{N}$ (1.6)

$E(\pi_N) = \pi; \;\; D(\pi_N) = 16\frac{pq}{N} = 3\frac{\pi^2}{N}$, таким чином (1.6) дає оцінку. Як відомо, помилка обчислень за методом Монте-Карло зазвичай пропорційна $\sqrt{\frac{d}{N}}$, де d - деяка константа, а N - кількість випробувань. З формули очевидно, що для підвищення точності в 10 разів необхідно збільшити кількість випробувань в 100 разів, а це означає, що метод Монте-Карло вимагає великих обчислювальних ресурсів[3].

Точність обчислень методу дуже сильно залежить від якості використовуваного генератора псевдовипадкових чисел, швидкість обчислень визначається функцією, яка описує аналізований процес і, звичайно ж, продуктивністю самого «обчислювача», що вже зазначалося вище. Спрощено схему алгоритму можна представити у вигляді, показаному на рис. 1.3[4].

Рисунок 1.3 Схема обчислень за методом Монте-Карло

Необхідно вивчити можливість розпаралелювання обчислень, заснованих на Методі Монте-Карло.

Дослідження

Якщо проаналізувати структуру, то можна прийти до висновку що проміжні розрахунки в методі Монте-Карло можуть здійснюватися незалежно на різних, окремо взятих «обчислювачах», а результати вже оброблятися на якомусь теж окремо взятому «аналізаторі» (рис. 1.4).

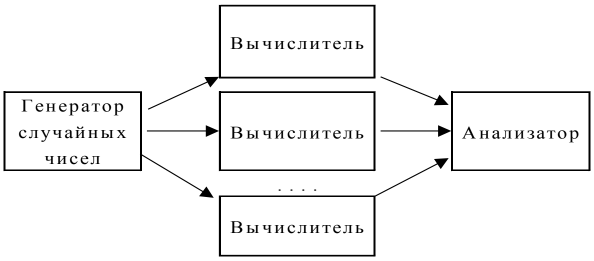

Рисунок 1.4 - Спрощений алгоритм паралельного обчислення за методом Монте-Карло

Згідно рис. 1.4, генератор випадкових чисел один для системи і, «проходячи» зверху вниз по «обчислювачам», «видає» кожному з «обчислювачів» генероване випадкове число. Ось тут криється підводний камінь, що знижує продуктивність даної паралельної системи. Інформація повинна бути передана між комп'ютерами[5].

Але, якщо числа випадкові, існує ймовірність і досить висока того, що серед їх послідовності будуть виникати повторення або дуже близькі значення, що не вплине на (нашу допустиму похибку обчислень). Отже, з'являється можливість дещо видозмінити саму схему обчислень (рис. 1.5) , надавши кожному «обчислювача» окремий, свій власний генератор випадкових чисел.

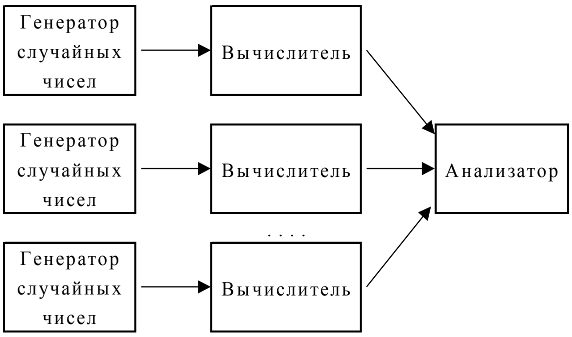

Рисунок 1.5 - Видозмінений алгоритм паралельного обчислення за методом Монте-Карло

Ця модель може бути реалізована на відкритих, так як їх архітектура спочатку орієнтована на велику кількість потоків, кожен з яких буде містити функції генератора випадкових чисел і обчислювача.

При цьому ця реалізація не буде містити недоліку - передачі великої кількості даних між комп'ютерами, так як кожен потік обчислень буде відносно незалежний від інших за рахунок власного генератора псевдовипадкових чисел. Але для реалізації цього необхідно розробити відповідний алгоритм генерації псевдовипадкових чисел.

ВИСНОВОК

Провівши аналіз областей застосування методів Монте-Карло, а також основних принципів організації обчислень цим методом, я прийшов до висновку що вивчення даного завдання актуально, і необхідно для підвищення ефективності обчислень поставлених завдань. Найбільш перспективним напрямком організації обчислень цим методом - є розпаралелення обчислень на відеокартах, що дозволить отримати зростання продуктивності обчислень, і підвищення його точності, за рахунок проведення більшої кількості випробувань.

Найбільш вузьке місце є генератори псевдовипадкових чисел, так як стандартні методи генерації псевдовипадкових послідовностей не підходять для використання при обчисленнях на відео картах, з за свою громіздкість.

Необхідно вивчити розробити алгоритм генерації псевдовипадкових чисел, що буде володіти великим періодом повтору, а також не є ресурсоємним, для ефективного використання його при обчисленнях методами Монте-Карло на відеокартах. Також алгоритм повинен відрізнятися від звичайних генераторів псевдовипадкових послідовностей використанням даних про фізичний стан відеокарти, що носить теж випадковий характер.

ЛІТЕРАТУРА

1. Gamerman, D. Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference. Boca Raton, FL: CRC Press, 1997.
2. Gentle J. Random Number Generation and Monte Carlo Methods. Springer-Verlag NY, 1998.
[1] 3. Gilks, W. R.; Richardson, S.; and Spiegelhalter, D. J. (Eds.). Markov Chain Monte Carlo in Practice. Boca Raton, FL: Chapman & Hall, 1996.
4. Gourdon X., Sebah P. and its computation through the ages. April 16, 2003. [http://numbers.computation.free.fr/Constants/constants.html]
5. Hoffman, P. The Man Who Loved Only Numbers: The Story of Paul Erdos and the Search for Mathematical Truth. New York: Hyperion, pp. 238-239, 1998.
6. Kuipers, L. and Niederreiter, H. Uniform Distribution of Sequences. New York: Wiley, 1974.