сообщения
объединенного
института
ядерных
исследований
дубна

P11-85-757

О.Н.Казаченко, Г.Г.Тахтамышев

ИСПОЛЬЗОВАНИЕ КВАЗИСЛУЧАЙНЫХ ЧИСЕЛ ПРИ МОДЕЛИРОВАНИИ

1985

ВВЕДЕНИЕ

Цель настоящей работы - дальнейшая проверка генераторов квазислучайных чисел. Первые результаты тестирования были опубликованы в [1,2]. В связи с широким использованием метода Монте-Карло и большими затратами времени ЭВМ на эти расчеты, разработка алгоритмов, позволяющих существенно ускорить получение результата, представляется нам важным направлением. В литературе /см., например,[11] с.254/ встречается утверждение, что при использовании квазислучайных чисел относительная погрешность результата убывает по закону, близкому к $1/N$ /где N - число точек/.

Поскольку традиционные генераторы псевдослучайных чисел способны обеспечить лишь сходимость $\sim 1/\sqrt{N}$, т.е. гораздо более медленную, понятно, что переход на генераторы квазислучайных чисел сулит существенный выигрыш в точности и в затратах времени ЭВМ. Однако практически переход на новые генераторы возможен только после проведения надежных серьезных тестов. Такой проверке посвящены как работы [1,2], так и настоящая работа.

1. МЕТОД ПРОВЕРКИ

Простым и естественным способом проверки качества генератора случайных чисел является вычисление многомерного интеграла, точное значение которого можно получить аналитическим путем. Для теста использовались две функции:

$$f_1(x_1,\dots,x_k) = \prod_j^k \exp(-b_j \cdot |x_j - a_j|)$$

и

$$f_2(x_1,\dots,x_k) = \prod_j^k 1/[1 + b_j^2 \cdot (x_j - a_j)^2].$$

Область определения функций - единичный K-мерный гиперкуб, т.е. $0 < x_j < 1$. Было проведено несколько вариантов расчета, которые отличались выбором параметров a_j и b_j. Величины a_j в каждом варианте разыгрывались равномерно в интервале от 0 до 1; таким образом, точка максимума функции находилась внутри области определения. Параметры b_j вычислялись по формуле $b_j = c \cdot j$, причем константа c подбиралась таким образом, чтобы величина интеграла находилась в пределах от $0{,}49 \cdot 10^{-3}$ до $0{,}51 \cdot 10^{-3}$.

Поскольку максимальное значение подынтегральной функции в обоих случаях равно 1, величина интеграла является мерой сингулярности функции.

В общем случае меру сингулярности или "коэффициент заполнения" можно определить как

$$\zeta = \frac{\langle f \rangle \cdot \omega}{f_{max} \cdot \Omega},$$

где $\langle f \rangle$ - среднее значение функции, f_{max} - ее максимальное значение, ω - объем K-мерной области определения, Ω - объем минимального параллелепипеда, в который можно заключить область определения.

Геометрический смысл вводимого понятия "коэффициент заполнения" ζ понятен: если заключить и функцию, и K-мерную область определения в K+1 -мерный параллелепипед наименьшего объема, то ζ будет равно отношению интеграла от функции к объему этого параллелепипеда. Другими словами, параметр ζ показывает, насколько заполняет функция /включая и ее область определения/ этот минимальный K+1 -мерный параллелепипед. При моделировании физических процессов эту величину часто называют эффективностью /см., например,[10] стр.361/, однако этот термин имеет достаточно много различных значений. Очевидно, что чем меньше "коэффициент заполнения" ζ, тем труднее получить оценку интеграла с заданной погрешностью ϵ. Точность оценки интеграла существенным образом может зависеть от сингулярности подынтегральной функции, поэтому при сравнении различных процедур интегрирования этот параметр полезно фиксировать. Таким образом, варианты расчета отличались друг от друга следующими параметрами:

1/ алгоритмом бросания точки в единичный гиперкуб,
2/ выбором функции,
3/ размерностью гиперкуба K /10, 20, 30, 40/,
4/ выбором параметров a_j, b_j.

Если в K-мерный единичный гиперкуб брошено равномерно n точек и в каждой точке вычислено значение функции f_i, можно получить оценку интеграла от функции f:

$$I_n = \frac{1}{n}\sum_{i=1}^{n} f_i . \qquad /1/$$

Относительная погрешность оценки

$$\epsilon_n = (I_n - I_0)/I_0, \qquad /2/$$

где I_0 - точное значение интеграла, полученное аналитическим путем.

Имея m независимых серий /каждая из которых состоит из n точек/, можно вычислить среднеквадратичную относительную погрешность

$$\langle \epsilon_n \rangle = \sqrt{\frac{1}{m}\sum_{\ell=1}^{m} (\epsilon_{n\ell})^2} \qquad /3/$$

и оценить ошибку этой величины[1,3]:

$$\Delta\langle \epsilon_n \rangle = \sqrt{\frac{1}{m}\Big[\frac{1}{m}\sum_{\ell=1}^{m} \epsilon_{n\ell}^4 - \Big(\frac{1}{m}\sum \epsilon_{n\ell}^2\Big)^2\Big]}. \qquad /4/$$

Обычно в каждом варианте расчета общее количество точек было равно $2048 \cdot 10^3$ /т.е. немного больше, чем 2 млн./. В процессе расчета делались вычисления, которые позволяли затем разбить эту последовательность на m серий, каждая длиной n точек /$m \cdot n = 2048000$/. Длина серии n варьировалась от 2000 до 1024000, а число серий m, соответственно, от 1024 до 2.

2. ГЕНЕРАТОРЫ КВАЗИСЛУЧАЙНЫХ ЧИСЕЛ

По-видимому, первый способ получения чисел, которые сейчас получили название квазислучайных, был предложен Н.М.Коробовым[4]. Коробовский алгоритм вычисления многомерных интегралов хорошо известен и включен в библиотеки стандартных программ ЭВМ всех крупных научных центров. Однако, поскольку в этом методе длина серии n не есть произвольное число и серия не может повторяться многократно, его тестирование способом, используемым в данной работе не может быть проведено.

Мы проводили проверку алгоритмов Соболя, Холтона и Рихтмайера и их сравнение с генератором псевдослучайных чисел. В качестве псевдослучайного генератора использовалась стандартная подпрограмма-функция RNDM, включенная в библиотеку стандартных программ ЭВМ серии ЕС. Последовательность Соболя реализуется так. Чтобы получить координаты i-й точки $(Q_1^i, Q_2^i, \dots, Q_k^i)$, нужно представить число i в двоичном виде: $i = e_p e_{p-1} \dots e_2 e_1$, а затем вычислить

$$Q_j^i = e_1 V_{1j} * e V_{2j} * \dots * e_p V_{pj} ,$$

где символ $*$ означает операцию поразрядного сложения по модулю 2 в двоичной системе. В работе[6] приведена таблица чисел V, с помощью которой можно получать до $2 \cdot 10^8$ точек в гиперкубе размерности не выше 51. Указан также способ расширения этой таблицы для получения большего числа точек.

Метод Холтона[7], известный также как метод Ван дер Корпута, состоит в следующем. Каждому измерению единичного куба j ставится в соответствие целое число r_j, причем все r_j взаимно простые. Для получения j-й компоненты i-й случайной точки (Q_j^i) число i представляется в r_j-ичной системе счисления: $i = e_p e_{p-1} \dots e_2 e_1$, и искомая компонента представляется как $Q_j^i = 0, e_1 e_2 \dots e_p$ (опять-таки в r_j-ичной системе).

Наиболее простым является алгоритм Рихтмайера[8]. Как и в предыдущем случае, каждому измерению куба j ставится в соответствие целое число S_j и все S_j опять взаимно простые. Затем Q_j^i вычисляются как $Q_j^i = \{i \cdot \sqrt{S_j}\}$, где $\{z\}$ означает дробную часть числа z.

3. РЕЗУЛЬТАТЫ И ПЕРСПЕКТИВЫ

Как было сказано выше, тестирование проводилось для трех квазислучайных генераторов /Соболь, Холтон, Рихтмайер/ и одного псевдослучайного /датчик RNDM/ на функциях двух типов с варьированием параметров /но с фиксацией в определенных пределах величины интеграла/ и для размерностей гиперкуба 10, 20, 30 и 40. Общее число проделанных тестов свыше 50. Типичная длина последовательности для одного варианта составляла 2048000 точек. Расчеты проводились на ЭВМ ЕС-1040 ЛЯП ОИЯИ.

Основная характеристика, которая нас интересовала - это зависимость среднеквадратичной относительной погрешности /формулы /3/ и /4// от длины серии n.

На рисунке показаны некоторые из полученных результатов.

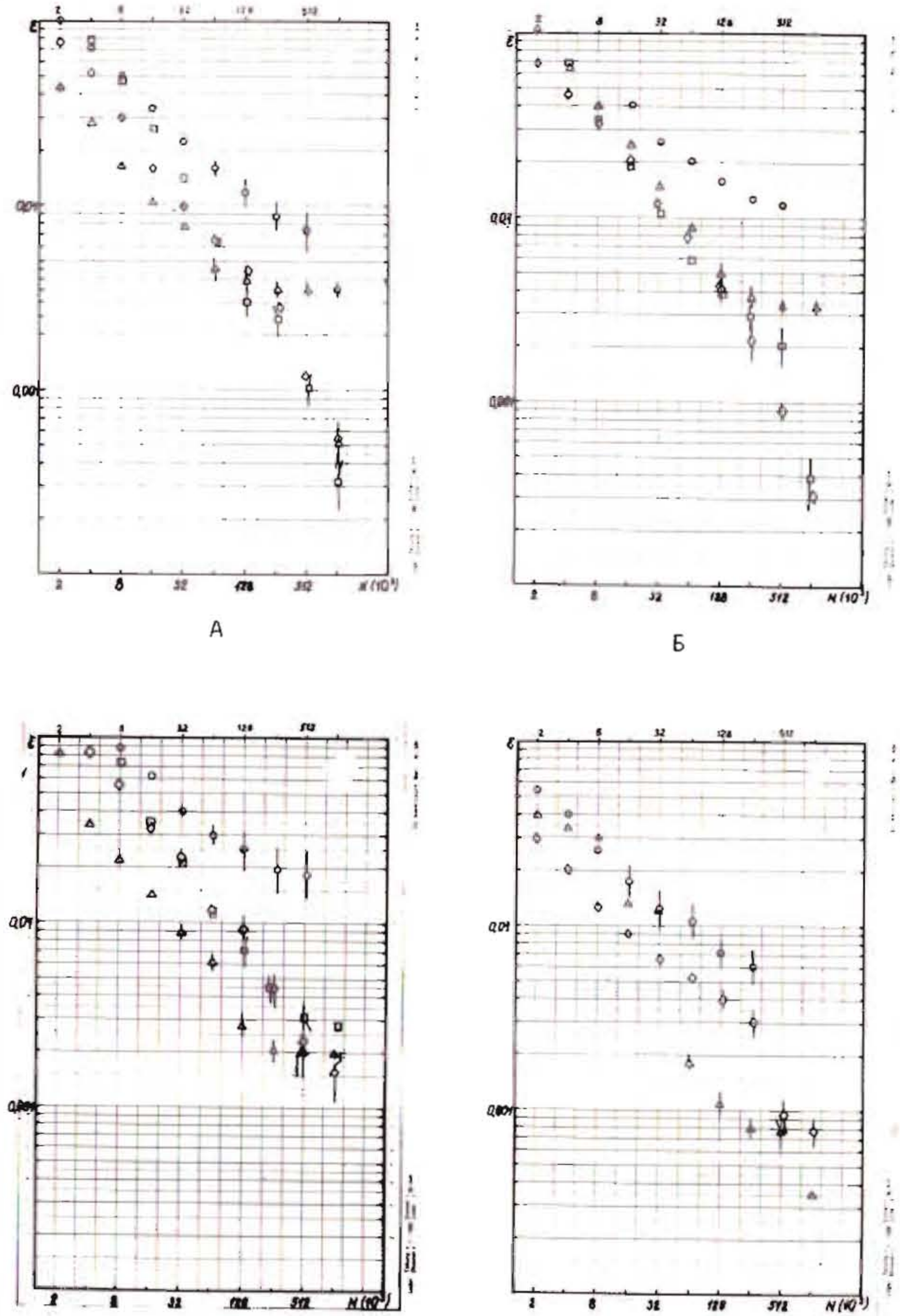
Анализ полученных закономерностей позволяет сделать два следующих основных вывода:

1. Все три типа квазислучайных генераторов во всех проведенных тестах обеспечили существенно лучшую точность расчета, чем псевдослучайный генератор.

2. Относительная погрешность расчета $\langle\epsilon\rangle$ для квазислучайных генераторов с увеличением длины серии n уменьшается быстрее, чем $1/\sqrt{n}$, но медленнее, чем $1/n$.

Эти выводы, однако, не следует трактовать как неоспоримое доказательство преимущества квазислучайных генераторов; скорее их следует рассматривать как доводы в пользу дальнейшего проведения тестов. Видно, что дальнейшие тесты необходимы и для более детального сравнения квазислучайных генераторов между собой.

По-видимому, следует расширить класс функций, на которых проводится тестирование. В этой связи следует отметить недавно появившуюся работу[9], в которой обсуждаются методы тестирования алгоритмов вычисления многомерных интегралов. В этой работе предлагается проводить тестирование на функциях шести типов; два из



А Б В Г

Относительная погрешность вычисления интеграла по методу Монте-Карло в зависимости от числа точек. Размерность области интегрирования для случаев А, Б и В равна 10, для случая Г - 40. Обозначения: о - генератор псевдослучайных чисел RNDM; Δ - генератор Соболя; ◇ - генератор Рихтмайера; □ - генератор Холтона.

них совпадают с использованными в нашей работе. Общим является и утверждение о существенной зависимости результата тестирования от величины интеграла; правда, в роли количественного параметра в [9] выступает не коэффициент заполнения ζ, а другой, менее наглядный, параметр "трудности".

В нашей работе мы умышленно обошли молчанием вопрос о сравнительной скорости работы генераторов. С нашей точки зрения, этот вопрос не носит принципиального характера по следующим причинам:

1. Эта скорость может существенно меняться в зависимости, например, от искусства программиста и языка, на котором реализован алгоритм.

2. Эти скорости /даже относительные/ разные для разных ЭВМ, и, наконец, самое существенное:

3. Если квазислучайные алгоритмы на самом деле окажутся такими хорошими, какими они сейчас кажутся, не представит большого труда реализовать их аппаратным путем /т.е. в виде электронного блока, соединенного с ЭВМ/. Это позволит практически свести к нулю время на выработку случайного числа.

ЛИТЕРАТУРА

1. Белов В.А., Клименко С.В., Тахтамышев Г.Г. Препринт ИФВЭ, 82-11, Серпухов, 1982.
2. Тахтамышев Г.Г. Препринт ИФВЭ АН КазССР, 83-06, Алма-Ата, 1983.
3. Корн Г., Корн Т. Справочник по математике. "Наука", М., 1968, с.539.
4. Коробов Н.М. ДАН СССР, 1959, 124, с.1207; Коробов Н.М. Теоретико-числовые методы в приближенном анализе. Физматгиз, М., 1963.
5. Соболь И.М., Статников Р.Б. Многомерные квадратурные формулы и функции Хаара. Физматгиз, М., 1969.
6. Соболь И.М., Статников Р.Б. Выбор оптимальных параметров в задачах со многими критериями. "Наука", М., 1981.
7. Halton J.H. Num.Math., 1960, 2, p.84.
8. James F. CERN Preprint 68-15, Geneva, 1968.
9. Genz A. Int.Conf. Tools, Methods and Languages for Scientific and Engineering Computations. Paris, May 17-19, 1983, p.81.
10. Копылов Г.И. Основы кинематики резонансов. "Наука", М., 1970.
11. Соболь И.М. Численные методы Монте-Карло. "Наука", М., 1973.

Рукопись поступила в издательский отдел
23 октября 1985 года.