

УДК 519.6

ВЫЧИСЛЕНИЕ МНОГОМЕРНЫХ КУБАТУР НА ПОСЛЕДОВАТЕЛЬНОСТЯХ СОБОЛЯ

© 2022 г. А. А. Белов^{1, 2, *}, М. А. Тинтул¹

¹Федеральное государственное бюджетное образовательное учреждение высшего образования “Московский государственный университет имени М.В. Ломоносова”, физический факультет, Москва, Россия

²Федеральное государственное автономное образовательное учреждение высшего образования “Российский университет дружбы народов”, Москва, Россия

*E-mail: aa.belov@physics.msu.ru

Поступила в редакцию 14.02.2022 г.

После доработки 28.02.2022 г.

Принята к публикации 23.03.2022 г.

Многомерные интегралы возникают во многих областях физики. Для них наиболее работоспособны методы Монте-Карло на квазислучайных точках Соболя. Усовершенствован алгоритм Соболя, что привело к улучшению равномерности распределения точек в многомерном кубе. Предложена многосеточная стратегия расчета, позволяющая найти кубатуру одновременно со статистической оценкой ее точности, что существенно повышает надежность расчета.

DOI: 10.31857/S0367676522070080

ВВЕДЕНИЕ

Интегралы от функций многих переменных встречаются во многих областях физики. Приведем примеры. Перенос нейтронов, фотонов и других частиц в среде описывается уравнением для функции распределения; эта функция зависит от трех координат среды и трех компонент вектора скорости частицы, то есть число переменных равно шести. Для определения коэффициентов теплопроводности или электропроводности среды приходится вычислять интегралы рассеяния; в них входят компоненты векторов скоростей двух частиц до момента столкновения и после момента столкновения. Общее число переменных в таком интеграле равно двенадцати. Возникают задачи и с существенно большим числом переменных.

В простейшей постановке рассматривают вычисление интеграла в единичном p -мерном кубе V . Пусть $\mathbf{x} = (x_1, x_2, \dots, x_p)$ есть p -мерный вектор. Требуется найти

$$I = \int_V f(\mathbf{x}) d\mathbf{x} = \int_0^1 \dots \int_0^1 f(x_1, x_2, \dots, x_p) dx_1 dx_2 \dots dx_p. \quad (1)$$

Точность сеточных методов стремительно падает с увеличением размерности p . Для получения приемлемой точности приходится брать все больше и больше точек, что делает расчеты непомерно трудоемкими и очень затратными по времени. В связи с этим при больших размерностях ($p > 3$)

используется метод Монте-Карло. Он предполагает использование случайных чисел, которые являются математической абстракцией. На практике же приходится использовать последовательности, лишь имитирующие случайные числа. От выбора такой последовательности сильно зависит работоспособность метода.

Расчеты представительных тестовых интегралов показывают, что для получения хорошей точности важна не столько случайность, сколько равномерность распределения точек. Наиболее эффективными оказываются последовательности Соболя с так называемыми “магическими” числами точек $N = 2^n$, $n = 0, 1, \dots$

В данной работе получены следующие результаты. Во-первых, предложены смещенные точки Соболя – модификация, улучшающая равномерность распределения точек и повышающая точность кубатур. Во-вторых, построена многосеточная стратегия, дающая апостериорную оценку точности. Преимущества предлагаемых алгоритмов проиллюстрированы на представительных тестовых примерах.

ПСЕВДОСЛУЧАЙНЫЕ ТОЧКИ

Метод Монте-Карло основан на том, что в кубе V выбирают N случайных точек \mathbf{x}_j ; при этом число N может быть произвольным, в отличие от кубатурных формул на регулярных сетках. Куба-

турная формула имеет схожий вид с формулой средних. Однако оценка ее погрешности оказывается кардинально иной.

$$I_N = \frac{1}{N} \sum_{j=1}^N f(\mathbf{x}_j), \quad \Delta_N \sim \sqrt{Df N^{-1/2}},$$

$$Df = \int_V f^2(\mathbf{x}) d\mathbf{x} - \left[\int_V f(\mathbf{x}) d\mathbf{x} \right]^2. \quad (2)$$

Здесь Df есть дисперсия. Оценка погрешности имеет не мажорантный, а вероятностный характер: величина погрешности распределена по закону Гаусса с указанным в формуле стандартом. Напомним, что погрешность не превышает 1 стандарта с вероятностью 0.68.

Оценка погрешности (2) не зависит от размерности p . Случайные точки сильно проигрывают регулярным сеткам при $p = 1$ или $p = 2$. Уже при $p = 4$ зависимость погрешности от N для случайных точек и регулярных сеток одинакова. При дальнейшем увеличении размерности случайные точки оказываются более выгодными; выигрыш быстро увеличивается при возрастании p .

Формулы (2) предполагают, что случайные точки \mathbf{x}_j имеют равномерную плотность распределения в кубе V и не коррелированы. Однако строгих математических способов построения таких точек не найдено. Предложен ряд математических алгоритмов; получаемые при этом точки называют псевдослучайными. Построению псевдослучайных точек посвящена обширная литература, например, [1–13]. В отечественной и зарубежной литературе наиболее распространены следующие генераторы:

- вихрь Мерсенна (Mersenne twister) и быстрый вихрь Мерсенна (SIMD-oriented fast Mersenne twister);
- мультипликативный конгруэнтный генератор (Multiplicative congruential generator);
- мультипликативный генератор Фибоначчи с запаздыванием (64-bit multiplicative lagged Fibonacci generator);
- комбинированный множественный рекурсивный генератор (combined multiple recursive generator);
- генератор Philo4x32;
- генератор Threefry4x64;
- генератор Марсалья (Marsaglia's SHR3 shift-register generator);
- модифицированный генератор Subtract-with-Borrow (modified Subtract-with-Borrow generator);
- модифицированная последовательность Лемера.

Именно эти генераторы реализованы во многих коммерческих пакетах (например, Matlab).

Качество каждой последовательности псевдослучайных чисел проверяют с помощью некоторых наборов тестов, основанных на теории вероятности [14–17]. Но никакой набор тестов не может быть полным и исчерпывающим. Поэтому такие проверки носят ограниченный характер. Даже простейшие визуальные тесты показывают, что широко распространенные последовательности не обеспечивают достаточно хорошей равномерности заполнения единичного квадрата [18, 19]. Вопрос о влиянии такой неравномерности на фактическую погрешность кубатур остается недостаточно ясным.

ТОЧКИ СОБОЛЯ

Последовательность Соболя определена неоднозначно. В ней использованы так называемые направляющие числа, вид которых может варьироваться. В ранних работах [1] таблицы направляющих чисел были построены для размерностей $p \leq 13$ и чисел $n \leq 20$ ($N \leq 2^{20}$). Позднее были построены числа для больших p и N [20]. Правда, при этом направляющие числа также менялись. В настоящее время доступна программа [21]. Вариант открытого доступа содержит $p \leq 50$ и $n \leq 31$ ($N \approx 2 \cdot 10^9$). Коммерческий вариант программы имеет $p \leq 2^{16} - 1$.

Напомним, что последовательности Соболя отдельно строятся для каждого p . Нельзя из p -мерной последовательности Соболя получить последовательность меньшего числа измерений. То же относится к магическим отрезкам последовательностей Соболя.

Сама кубатурная формула Соболя имеет тот же вид, что и (2). Но оценка ее погрешности не вполне ясна. Распределение точек лишь при магических N приближается по свойствам к равномерному. При промежуточных N оно получается выбрасыванием части точек и теряет свойство равномерности. Поэтому для кубатур следует использовать лишь магические N .

Предпринимались различные попытки обобщить последовательности Соболя. Однако поиски оптимальных вариантов таких обобщений неизменно приводили опять к последовательностям Соболя. Поэтому мы относимся к таким обобщениям осторожно.

СМЕЩЕНИЕ ТОЧЕК СОБОЛЯ

Расположение точек Соболя несколько асимметрично. Например, если взять $N = 2^n$, то среднее арифметическое всех проекций точек на лю-

бую ось будет не 0.5, а $0.5(1 - 1/N)$. Очевидно, эта асимметричность не благоприятна для получения хорошей точности кубатур.

На рис. 1 черными кружками показаны двумерные точки Соболя для первых магических чисел. При $n = 0$ единственная точка лежит в углу единичного квадрата. Вычисление кубатуры по ней дает формулу первого порядка точности. Однако если сдвинуть эту точку на 0.5 по каждой координате, то кубатура по смещенной точке (светлый кружок) будет иметь второй порядок точности. Для случая $n = 1$ две точки расположены одна в углу квадрата и одна в центре, что также даст первый порядок точности. Но если сместить их на 0.25 по каждой координате, то погрешность кубатуры явно уменьшится. Поэтому можно предложить общий принцип смещения для любого числа измерений: если $N = 2^n$, то прибавим ко всем координатам всех точек $(2N)^{-1}$. Это смещение целесообразно применять только для магических чисел Соболя. При этом смещения различны для различных N .

МНОГОСЕТОЧНЫЙ РАСЧЕТ

Тестовые расчеты показывают, что фактическая погрешность Δ_N убывает как $O(N^{-1})$. Это наводит на мысль о возможности аппроксимации интеграла (и следовательно, его погрешности) как функции N . Однако эта аппроксимация не может быть гладкой, типа интерполяционной аппроксимации Ричардсона для сеточных методов. В данном случае точки получают статистическими методами, поэтому их обработку нужно вести с помощью среднеквадратичной аппроксимации. Для этого нужно выбрать вид аппроксимации и приписать точкам какие-то веса.

В качестве рабочей гипотезы мы предположили закон убывания погрешности $\Delta_N \sim N^{-1}$. Но поскольку при увеличении p явно проявляется статистический характер погрешности, то стандартное уклонение этих погрешностей мы приняли пропорциональными в степени $N^{-1/2}$. Это и есть вес, используемый при аппроксимации.

Была предложена следующая многосеточная процедура. Выполним расчет с магическими $N = 2^n$, $n = 10, 11, \dots$. В результате получим последовательность значений интеграла $\{I_N\}$. Теперь аппроксимируем эту последовательность по методу наименьших квадратов

$$I_N \approx a + bN^{-1} \tag{3}$$

Здесь a – уточненное значение интеграла. Одновременно вычислим стандартное уклонение σ_a для величины a . Это стандартное уклонение яв-

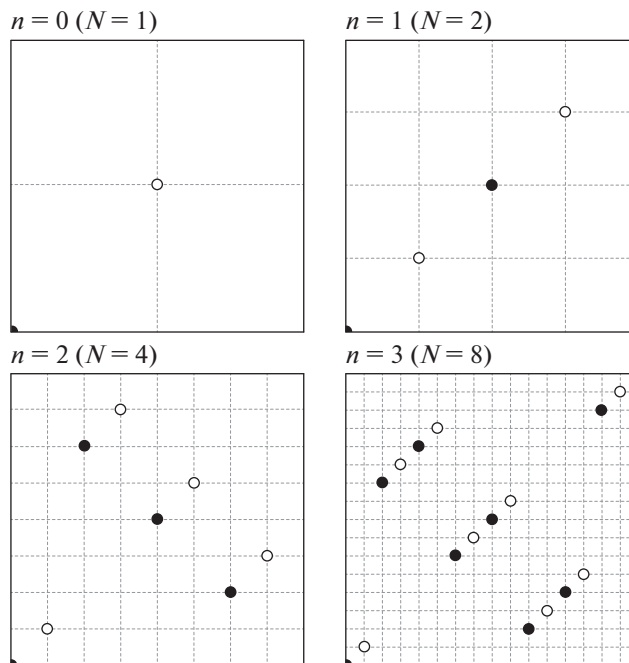


Рис. 1. Магические точки Соболя для $p = 2$: точки – несмещенные, кружки – смещенные; значения n указаны около квадратов.

ляется статистической оценкой точности для найденного значения интеграла.

Отметим, что начало последовательности $\{I_N\}$, соответствующее $n = 0, 1, \dots, 9$ не учитывается при аппроксимации (3), поскольку эти сетки недостаточно подробны, и скорость убывания погрешности еще не соответствует $O(N^{-1})$.

ТЕСТОВЫЙ ИНТЕГРАЛ

Численные эксперименты целесообразно проводить на многомерных интегралах по единичному кубу, точные значения которых известны. Тогда можно непосредственно определять погрешность численного расчета и исследовать ее поведение. Обсудим, какие требования целесообразно предъявлять к подынтегральной функции.

В многомерных задачах используют понятие эффективной размерности функции. Например, рассмотрим две функции:

$$f(\mathbf{x}) = \prod_{j=1}^p f_j(x_j) \tag{4}$$

и

$$f(\mathbf{x}) = f_1\left(\sum_{j=1}^p \alpha_j x_j\right), \tag{5}$$

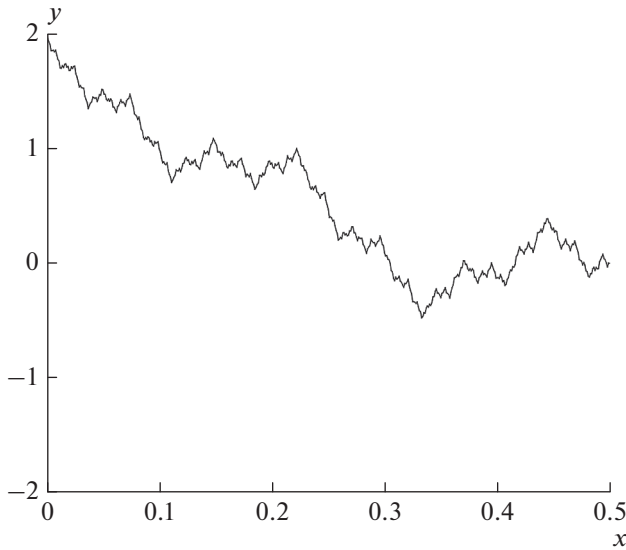


Рис. 2. Функция Вейерштрасса при $a = 3, b = 0.5$.

где все $f_j(x_j)$ существенно отличны от констант. В первой функции все переменные в равной степени важны, и эффективная размерность функции равна p . Вторая функция зависит лишь от одной комбинации переменных, так что ее эффективная размерность равна 1. Чем больше эффективная размерность функции, тем задача труднее. Поэтому наиболее трудными оказываются функции первого типа.

Заметим, что не нарушая общности, можно считать все подынтегральные функции неотрицательными: их можно сделать неотрицательными путем прибавления константы, а такая процедура не влияет на погрешность любого метода. Пусть для функции типа произведения каждая f_j существенно отлична от нуля лишь на участке длиной β своего единичного ребра. Тогда произведение одномерных функций будет значимо отличаться от нуля в объеме β^p . Если β невелико, то при увеличении p объем β^p стремительно убывает; например, при $\beta = 0.1$ и $p = 10$ величина $\beta^p = 10^{-10}$. В этом случае для получения приемлемой точности любой метод Монте-Карло потребует числа узлов $N \gg \beta^{-p}$. Видно, что для того, чтобы число точек было разумным, следует брать β близким к единице.

С учетом указанных соображений, в качестве теста мы выбрали тест вида (4). Он не является легким, несмотря на внешнюю простоту. Все f_j будем считать одинаковыми и равными функции Вейерштрасса

$$f_j(x_j) = \sum_{n=0}^{\infty} b^n \cos(a^n \pi x_j), \quad (6)$$

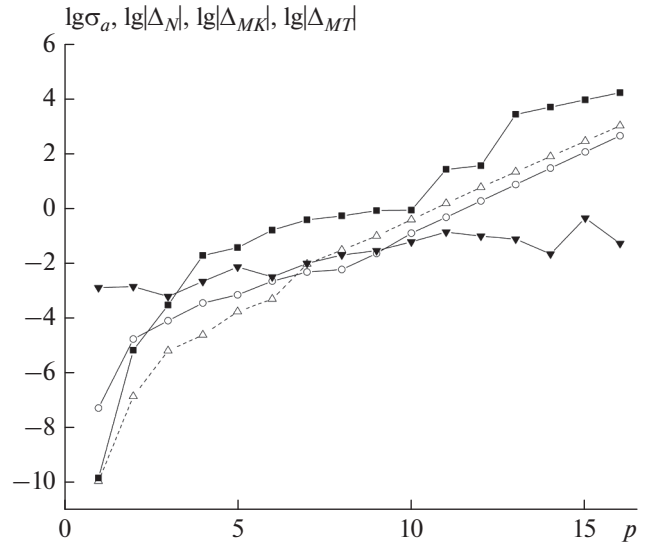


Рис. 3. Логарифм относительной погрешности расчета интеграла от функции Вейерштрасса для $N = 2^{20}$: светлый треугольник — Δ_N , кружок — σ_a для смещенных точек Соболя, черный перевернутый треугольник — вихрь Мерсенна, черный квадратик — метод трапеций.

где a — произвольное нечетное число, не равное единице, а b — положительное число, меньшее единицы. Известно, что при условиях $ab \geq 1, a > 1$ функция Вейерштрасса непрерывна, но не имеет производной ни в одной точке. Такой тест исключительно сложен. Вид функции Вейерштрасса показан на рис. 2. С учетом симметрии функции Вейерштрасса интегрирование будем проводить по кубу со сторонами $x_j \in [0, 0.5]$. Для удобства отнормируем функцию Вейерштрасса так, чтобы выполнялось условие нормировки

$$\int_V f(x) dx = 1. \quad (7)$$

РЕЗУЛЬТАТЫ РАСЧЕТОВ

Подсчитаем интеграл от многомерной функции Вейерштрасса (4), (6) с помощью трех качественно разных подходов: регулярной кубатуры на формулах трапеций, классического метода Монте-Карло с использованием Вихря Мерсенна и смещенных магических точек Соболя.

Сравним эти три подхода по величине погрешности при достаточно скромном числе точек $N = 2^{20}$. На рис. 3 приведены логарифмы погрешностей в зависимости от размерности p . Проанализируем приведенные кривые.

Вихрь Мерсенна

Начиная с размерности $p = 11$, ниже всех лежит кривая, соответствующая вихрю Мерсенна. Несмотря на хорошую точность, у нас нет никаких средств подтвердить ее. Попытка применить к вихрю Мерсенна среднеквадратичную аппроксимацию (3) не увенчалась успехом: величины σ_a оказываются то больше, то меньше фактической погрешности в зависимости от размерности p , причем отличие может быть существенным.

Оценкой погрешности вихря Мерсенна может служить стандарт $(Df/N)^{1/2}$, однако расчет дисперсии для некоторых интегралов может оказаться проблематичным. Кроме того, работоспособность метода Монте-Карло сильно зависит от выбора последовательности, имитирующей случайные числа, поэтому стандарт и фактическая погрешность могут сильно различаться.

В целом, величина $\lg|\Delta_{МК}|$ лежит в пределах от -3.5 до 0 и медленно растет с увеличением размерности p .

Формула трапеций

Ее погрешность определяется формулой

$$|\Delta_N| \leq \frac{1}{12k^2} \max \left| \frac{d^2 f_j}{dx^2} \right|. \quad (8)$$

где $k = N^{1/p}$ – число узлов по каждой координате. Тем самым, ее погрешность есть $O(N^{-2/p})$; точность должна быстро убывать с ростом p . Соответствующая кривая (маркер – черный квадрат на рис. 3) иллюстрирует неплохую точность $\lg|\Delta_{MT}| \approx -5.2$ при $p = 2$; это много точнее классического метода МК. Однако с увеличением p погрешность быстро возрастает, и уже при $p \geq 4$ превосходит погрешность метода МК. При еще больших размерностях метод трапеций быстро становится неконкурентоспособным.

Последовательность Соболя

Несмотря на то, что при больших размерностях p наилучший результат показывает вихрь Мерсенна, у смещенных точек Соболя есть разумная оценка точности – стандартное отклонение σ_a . Таким образом, даже в сложных задачах мы можем апостериорно оценить фактическую точность с помощью σ_a , увеличить число точек и повторить расчет. Это особенно важно для многомерных интегралов с неизвестным точным ответом.

Сравним поведение кривых для сдвинутых точек Соболя и для вихря Мерсенна. Точки Соболя

строятся неслучайным образом, и их расположение в многомерном кубе имеет некоторые закономерности. Например, заполнение единичного квадрата на плоскости имеет характер регулярного кружева [19]. При заполнении трехмерного куба точки ложатся на систему параллельных плоскостей [22], и т.д. С повышением размерности статистические свойства последовательностей Соболя (как исходных, так и сдвинутых) ухудшаются. Поэтому точность кубатур при фиксированном числе точек с ростом размерности также ухудшается (см. рис. 3). Если зафиксировать размерность интеграла и увеличивать число точек, то точность кубатур на последовательностях Соболя будет убывать как $O(1/N)$, причем коэффициент пропорциональности увеличивается с увеличением размерности.

Вихрь Мерсенна имеет примерно одинаковые статистические свойства при различных размерностях $1 \leq p \leq 16$ [14]. Поэтому данный генератор при фиксированном числе точек дает примерно одинаковую погрешность во всем диапазоне размерностей (см. рис. 3). При фиксированной размерности и увеличении N погрешность вихря Мерсенна убывает только как $O(N^{-1/2})$. Это существенно медленнее, чем для точек Соболя. Иными словами, с увеличением N кривая для вихря Мерсенна на рис. 3 будет довольно медленно сдвигаться вниз как единое целое.

Из сказанного следует, что найдется такая пара p_0 и N_0 , при которых погрешности кубатур на последовательностях Соболя и на вихре Мерсенна сравниваются. При $p < p_0$ и $N = N_0$ более точными оказываются последовательности Соболя, а при $p > p_0$ – вихрь Мерсенна. Значения p_0 и N_0 зависят от конкретной подынтегральной функции. Для рис. 3 они равны $N_0 = 2^{10}$, $p_0 = 6$.

ЗАКЛЮЧЕНИЕ

Для вычисления многомерных интегралов наиболее эффективны магические точки Соболя. Предложено усовершенствование этих последовательностей – смещенные магические точки Соболя, обеспечивающие большую равномерность распределения точек в многомерном кубе. Это ощутимо повышает точность кубатур.

Существенной трудностью методов Монте-Карло является апостериорное подтверждение фактической точности. В данной работе предложен многосеточный алгоритм, позволяющий найти сеточное значение интеграла одновременно со статистически достоверной оценкой его точности. Ранее такие оценки были неизвестны.

Проведены расчеты представительных тестовых интегралов с высокой фактической размер-

ностью p (до $p = 16$). Рассматривались гладкие подынтегральные функции, а также многомерная функция Вейерштрасса, не имеющая производной ни в одной точке. Эти расчеты убедительно показывают преимущества предложенных методов.

Работа поддержана выполнена при поддержке Совета по грантам Президента РФ (проект № МК-3630.2021.1.1).

СПИСОК ЛИТЕРАТУРЫ

1. *Соболь И.М.* Численные методы Монте-Карло. М.: Наука, 1973. 312 с.
2. *Кнут Д.* Искусство программирования. Т. 2. Полночисленные алгоритмы. М.: "Вильямс", 2007. 788 с.
3. *Fishman George S.* Monte Carlo: concepts, algorithms and applications. N.Y.: Springer, 1996. 698 p.
4. *Makoto Matsumoto, Takuji Nishimura* // ACM Trans. Model. Comput. Simul. 1998. V. 8. No. 1. P. 3.
5. *Takuji Nishimura* // ACM Trans. Model. Comput. Simul. 2000. V. 10. No. 4. P. 348.
6. <http://www.math.sci.hiroshima-u.ac.jp/m-mat/MT/emt.html>.
7. *Park S.K., Miller K.W.* // Commun. ACM. 1998. V. 31. No. 10. P. 1192.
8. *Mascagni M., Srinivasan A.* // Parallel Comput. 2004. V. 30. No. 7. P. 899.
9. *L'Ecuyer P.* // Oper. Res. 1999. V. 47. No. 1. P. 159.
10. *Salmon J.K., Moraes M.A., Dror R.O., Shaw D.E.* // Proc. Intern. Conf. High Perform. Comp. Network. Stor. Anal. (SC11). (N.Y., 2011).
11. *Marsaglia G., Tsang W.W.* // J. Stat. Software. 2000. V. 5. No. 8. P. 1.
12. *Marsaglia G., Zaman A.* // Ann. Appl. Prob. 1991. V. 1. No. 3. P. 462.
13. *Wichmann B.A., Hill I.D.* // Appl. Statistics. 1982. V. 31. No. 2. P. 188.
14. *Цветков Е.А.* // Матем. моделир. 2011. Т. 23. № 5. С. 81.
15. <https://web.archive.org/web/20160125103112/http://stat.fsu.edu/pub/diehard/>.
16. *L'Ecuyer P., Simard R.* // ACM Tran. Math. Software. 2007. V. 33. No. 4. Art. No. 22.
17. *Rukhin A., Soto J., Nechvatal. J.* A statistical test suite for random and pseudorandom number generators for cryptographic applications. NIST Special Publication. Art. No. 800-22.
18. *Белов А.А., Калиткин Н.Н., Тинтул М.А.* // Препринты ИПИМ им. М.В. Келдыша. 2019. № 137. 28 с.
19. *Белов А.А., Калиткин Н.Н., Тинтул М.А.* // ЖВММФ. 2020. Т. 60. № 11. С. 1807; *Belov A.A., Kalitkin N.N., Tintul M.A.* // Comput. Math. Math. Phys. 2020. V. 60. No. 11. P. 1747.
20. *Соболь И.М.* // ЖВММФ. 1976. Т. 16. № 5. С. 1332.
21. <https://www.broda.co.uk/software.html>.
22. *Калиткин Н.Н., Альшина Е.А.* Численные методы. Кн. 1. Численный анализ. М.: Академия, 2013. 304 с.

Calculation of multidimensional cubatures on Sobol sequences

A. A. Belov^{a, b, *}, M. A. Tintul^a

^a *Lomonosov Moscow State University, Moscow, 119991 Russia*

^b *Peoples' Friendship University of Russia (RUDN University), Moscow, 117198 Russia*

^{*}*e-mail: aa.belov@physics.msu.ru*

Multidimensional integrals arise in many areas of physics. Monte Carlo methods based on quasi-random Sobol points are the most efficient for their calculation. In this paper we propose an improvement to the Sobol algorithm which improves the uniformity of the distribution of points in a multidimensional cube. A multigrid calculation strategy is proposed that allows finding the cubature simultaneously with a statistical assessment of its accuracy. This significantly increases reliability of the calculation.