



ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 8 за 2021 г.



ISSN 2071-2898 (Print)
ISSN 2071-2901 (Online)

А.А. Белов, [Н.Н. Калиткин](#),
М.А. Тинтул

Многомерные кубатуры на
последовательностях
Соболя

Рекомендуемая форма библиографической ссылки: Белов А.А., Калиткин Н.Н., Тинтул М.А. Многомерные кубатуры на последовательностях Соболя // Препринты ИПМ им. М.В.Келдыша. 2021. № 8. 24 с. <https://doi.org/10.20948/prepr-2021-8>
<https://library.keldysh.ru/preprint.asp?id=2021-8>

**Ордена Ленина
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В.Келдыша
Российской академии наук**

А. А. Белов, Н. Н. Калиткин, М. А. Тинтул

**МНОГОМЕРНЫЕ КУБАТУРЫ
НА ПОСЛЕДОВАТЕЛЬНОСТЯХ СОБОЛЯ**

Москва — 2021

Белов А. А., Н. Н. Калиткин, Тинтул М. А.

Многомерные кубатуры на последовательностях Соболя

Вычисление многомерных кубатур в единичном кубе является сложной задачей численных методов, а ее прикладное значение велико. В работе сравниваются различные методы вычисления: произведение регулярных одномерных сеточных формул, классический метод Монте-Карло с псевдослучайными точками и последовательности Соболя. Предложено использовать не любые последовательности Соболя, а только с магическими числами $N = 2^n$. Кроме того, предложены смещенные точки Соболя: у магических точек Соболя все координаты одновременно увеличиваются на величину $(2N)^{-1}$. Сравнения на тесте показали, что последний способ существенно превосходит по точности все остальные.

Ключевые слова: многомерные кубатуры, метод Монте-Карло, последовательности Соболя, магические числа, смещенные точки Соболя

Aleksandr Aleksandrovich Belov, Nikolai Nikolaevich Kalitkin, Maksim Aleksandrovich Tintul

Multidimensional cubatures on Sobol sequences

Calculation of the multidimensional cubatures in the unit hypercube is a complex problem of numerical methods, and its application value is great. This paper compares various calculation methods: product of regular one-dimensional grid formulae, classical Monte Carlo method using pseudorandom points and Sobol sequences. It is suggested to use not any Sobol sequences, but only with magic numbers $N = 2^n$. In addition, offset Sobol points are proposed: all coordinates of magic Sobol points are simultaneously increased by $(2N)^{-1}$. Comparisons on the test showed that the latter method is significantly more accurate than all the others.

Key words: multidimensional cubatures, Monte Carlo method, Sobol sequences, magic numbers, offset Sobol points

1. Проблемы вычисления многомерных кубатур

1.1. Постановка задачи

Многомерные функции встречаются в разнообразных задачах физики. Приведем некоторые примеры. В задачах газодинамики, теории упругости и других все величины зависят от трех пространственных координат. 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

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

В физических задачах функция $f(\mathbf{x})$ обычно оказывается непрерывной и достаточно гладкой, то есть имеющей непрерывные производные нескольких порядков. Размерность p может быть достаточно большой, как мы видели выше. Возникает вопрос: какими методами вычислять интегралы (1)?

В математических постановках для некоторых приложений (но в основном для общности) рассматривают огромные размерности p порядка 100–1000, а также негладкие и разрывные подынтегральные функции. Однако здесь мы ограничимся случаем достаточно гладких функций и умеренных размерностей.

1.2. Сеточные методы

Интегралы малой размерности ($p \leq 3$) легко вычисляются по простейшим кубатурным формулам трапеций или средних на равномерных сетках. Особенно просто выглядит формула средних. Делят каждую сторону куба V на k равных частей. Тогда куб разбивается на $N = k^p$ кубиков со стороной $1/k$.

Обозначим центр j -го кубика как \mathbf{x}_j ; эти центры образуют регулярную сетку. Тогда приближенное значение интеграла и теоретическая оценка погрешности имеют следующий вид:

$$I_N = \frac{1}{N} \sum_{j=1}^N f(\mathbf{x}_j), \quad \Delta_N \equiv |I - I_N| \leq \frac{1}{24k^2} \sum_{i=1}^p \max \left| \frac{d^2 f}{dx_i^2} \right|. \quad (2)$$

Поскольку $k = N^{1/p}$, то погрешность имеет порядок величины $\Delta_N = O(N^{-2/p})$. Для одномерной задачи это дает эффективный порядок точности два, что считается достаточно хорошим для практических вычислений. Уже в двумерном случае эффективный порядок точности лишь первый. При дальнейшем увеличении p эффективный порядок точности стремительно уменьшается. Поэтому в многомерном кубе для получения удовлетворительной точности потребуется неприемлемо огромное число узлов N .

1.3. Псевдослучайные точки

Принципиально другой подход используется в методах Монте-Карло. В этом случае 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. Кубатурная формула записывается формально так же, как и формула средних. Однако оценка ее погрешности оказывается кардинально иной.

$$I_N = \frac{1}{N} \sum_{j=1}^N f(\mathbf{x}_j), \quad \Delta_N \sim \sqrt{Df N^{-1/2}}, \quad Df = \int_V f^2(\mathbf{x}) d\mathbf{x} - \left[\int_V f(\mathbf{x}) d\mathbf{x} \right]^2. \quad (3)$$

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 (3) 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 (3) предполагают, что случайные точки 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, 2].

Качество каждой последовательности псевдослучайных чисел проверяют с помощью некоторых наборов тестов, основанных на теории вероятности. 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 [3,4]. Вопрос о влиянии такой неравномерности на фактическую погрешность кубатур остается недостаточно ясным.

1.4. Точки Соболя

В 1973 году Соболев предложил точки, которые он назвал $ЛП_\tau$ – последовательностями; сейчас в зарубежной литературе их называют последовательностями Соболя. Эти последовательности не имитируют случайные распределения; наоборот, они стремятся не вполне случайным образом добиться равномерного распределения в многомерном кубе. Введем

Определение. Назовем магическими отрезки последовательности Соболя с числом точек, удовлетворяющих формуле

$$1 \leq j \leq N, N = 2^n (n = 0, 1, \dots - \text{целое число}). \quad (4)$$

Особенно просто описываются свойства последовательностей Соболя, если рассматривать только магические отрезки. При магическом N проекции всех точек на каждую ось принимают значения $j/N, 0 \leq j \leq N-1$. Такие проекции образуют равномерную сетку по каждой оси. Распределение проекций точек на двумерные плоскости также обладают определенными свойствами равномерности: если разбить двумерную грань на N одинаковых прямоугольников, то каждый прямоугольник содержит ровно одну точку последовательности. Это нетривиально, так как существует много разбиений на одинаковые прямоугольники. Например, при $n=2, N=4$ есть три типа разбиений: разбиение на 4 вертикальных прямоугольника размером $1/4 \times 1, 4$

горизонтальных прямоугольника размером $1 \times 1/4$ или 4 квадрата размером $1/2 \times 1/2$. При увеличении n количество типов разбиений стремительно возрастает; поэтому достижение равномерности в указанном смысле нетривиально. Сходные требования предъявляются к многомерным разбиениям.

Такие точки называют также квазислучайными. Сравнение квазислучайных и псевдослучайных точек показано на Рис. 1 для $N = 1024$, $n = 10$.

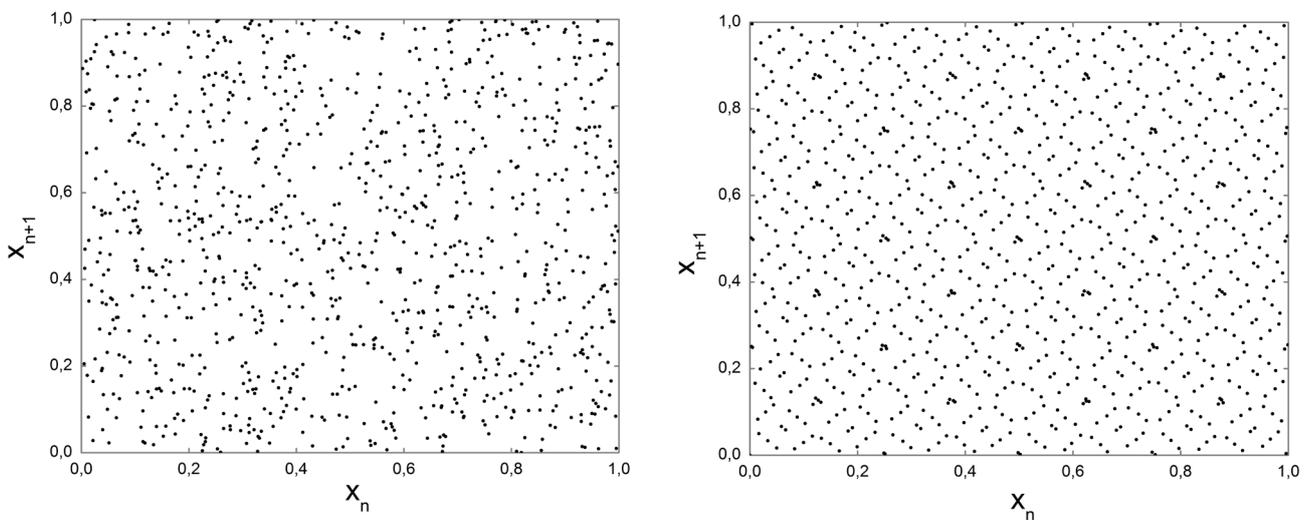


Рис.1. Точки для $p = 2$, $N = 1024$ ($n = 10$); слева – одна из реализаций Вихря Мерсенна, справа – последовательность Соболя.

На рис. 1 слева показаны точки одного из наиболее известных генераторов псевдослучайных чисел - Вихря Мерсенна; видно, что распределение весьма далеко от равномерного. На рис. 1 справа показаны точки Соболя; по специфическому рисунку видно, что распределение явно неслучайное, однако распределение гораздо более равномерно, чем для Вихря Мерсенна.

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 [5]. Правда, при этом направляющие числа также менялись. В настоящее время доступна программа [6]. 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 (3). 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 квазиравномерности. 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.

В данной работе мы с помощью численных экспериментов исследовали точность многомерных кубатур на последовательностях Соболя. При этом выяснялась фактическая точность полученных результатов и была произведена попытка повысить точность подобных кубатур. Совместная обработка серии расчетов с различным набором магических значений N позволила добиться точности заметно лучше, чем $O(N^{-1/2})$.

2. Исследование на тестах

2.1. Конструирование теста

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 подинтегральной функции.

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

$$f(x) = \prod_{j=1}^p f_j(x_j) \text{ и } f(x) = f_1\left(\sum_{j=1}^p \alpha_j x_j\right), \quad (5)$$

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.

Методы Монте-Карло применимы даже к разрывным функциям. Однако в значительной части физических приложений $f(\mathbf{x})$ является достаточно гладкой, то есть имеет несколько непрерывных производных. Поэтому при тестировании мы можем выбирать достаточно гладкие функции. С учетом указанных соображений мы выбрали несложный, но достаточно эффективный тест:

$$f(\mathbf{x}) = \prod_{j=1}^p \frac{\pi}{2} \sin \frac{\pi x_j}{2}, \int_V f(\mathbf{x}) d\mathbf{x} = 1, Df = \left(\frac{\pi^2}{8} \right)^p - 1. \quad (6)$$

Дисперсия достаточно медленно растет при увеличении p . Эффективная размерность теста равна p , то есть максимальна. Поэтому тест не является легким, несмотря на внешнюю простоту.

2.2. Визуализация погрешности

В данной работе мы ограничились размерностями пространства $p \leq 16$ и последовательностями Соболя длиной $N \leq 2^{20}$. Фиксировалась размерность p и

вычислялась кубатура тестовой функции (6) на всех магических отрезках последовательности Соболя с $N = 2^n, n = 0, 1, \dots, 20$. Для каждого отрезка вычитанием найденной кубатуры из единицы находилась погрешность Δ_N .

Погрешность Δ_N достаточно быстро убывала с увеличением N , и знак ее был неодинаков для разных N . Поэтому для визуализации погрешности нельзя было строить графики зависимости $\Delta(N)$ в двойном логарифмическом масштабе, как это делается для знакопостоянных погрешностей сеточных методов.

Как известно, в методах Монте-Карло со случайными точками погрешность кубатур убывает как $N^{-1/2}$, а ее поведение соответствует правилам статистики: погрешность может иметь любой знак и распределена по Гауссу со стандартом $(Df / N)^{1/2}$. Для начала мы приняли эту гипотезу и выбрали в качестве ординаты $(N / Df)^{1/2} \Delta_N$. Для каждого p были построены графики зависимости нормированной погрешности от аргумента $n = \log_2 N$. Они показаны на рис 2-14. Обсудим поведение погрешности на этих рисунках.

Если бы поведение погрешности соответствовало традиционным случайным точкам Монте-Карло, то нормированная погрешность беспорядочно меняла бы свой знак.

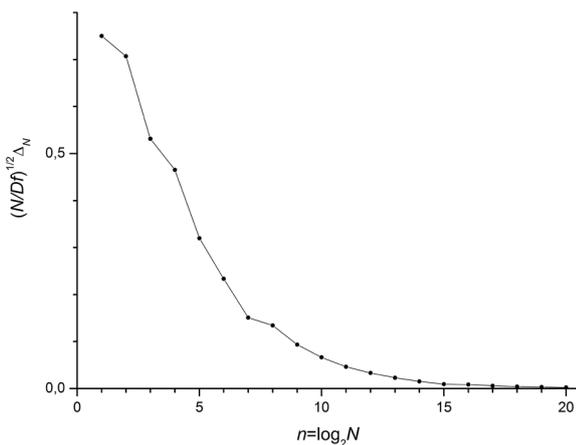


Рис.2. Зависимости нормированной погрешности $(N / Df)^{1/2} \Delta_N$ от $n = \log_2 N$ для последовательностей Соболя; $p = 2$.

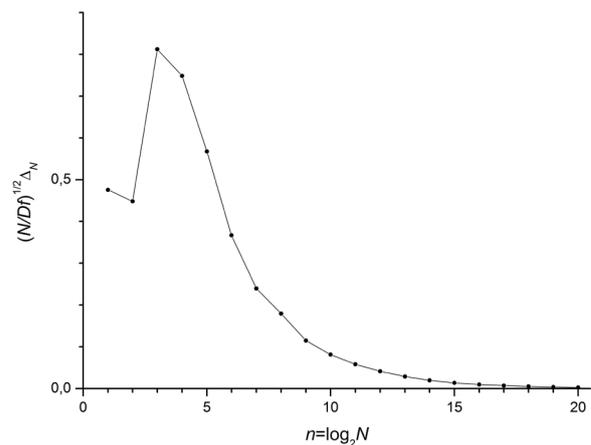


Рис.3. Зависимости нормированной погрешности $(N / Df)^{1/2} \Delta_N$ от $n = \log_2 N$ для последовательностей Соболя; $p = 3$.

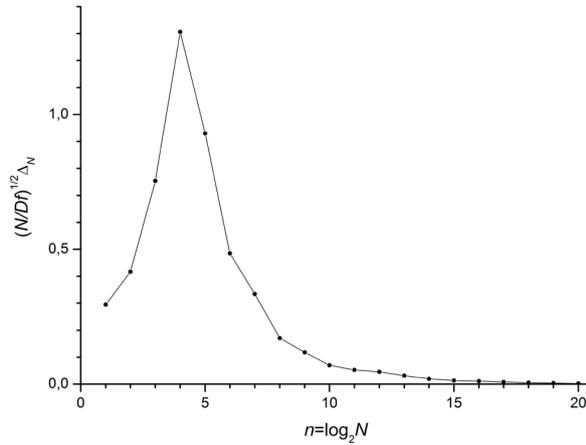


Рис.4. Зависимости нормированной погрешности $(N / Df)^{1/2} \Delta_N$ от $n = \log_2 N$ для последовательностей Соболя; $p = 4$.

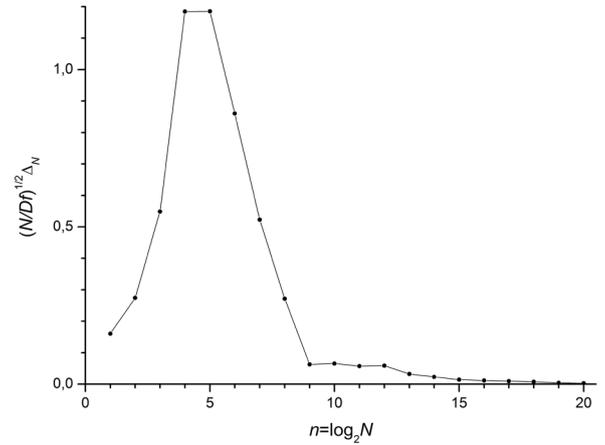


Рис.5. Зависимости нормированной погрешности $(N / Df)^{1/2} \Delta_N$ от $n = \log_2 N$ для последовательностей Соболя; $p = 5$.

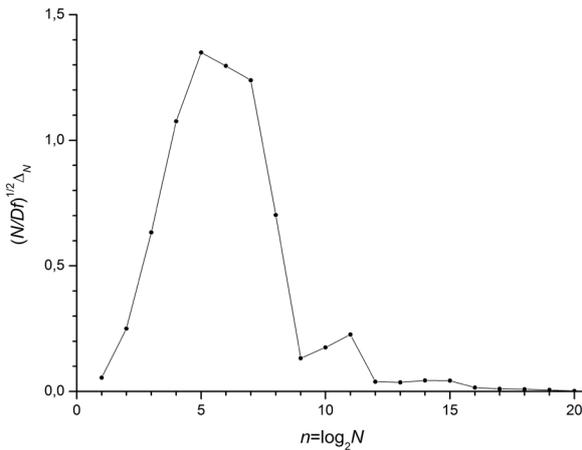


Рис.6. Зависимости нормированной погрешности $(N / Df)^{1/2} \Delta_N$ от $n = \log_2 N$ для последовательностей Соболя; $p = 6$.

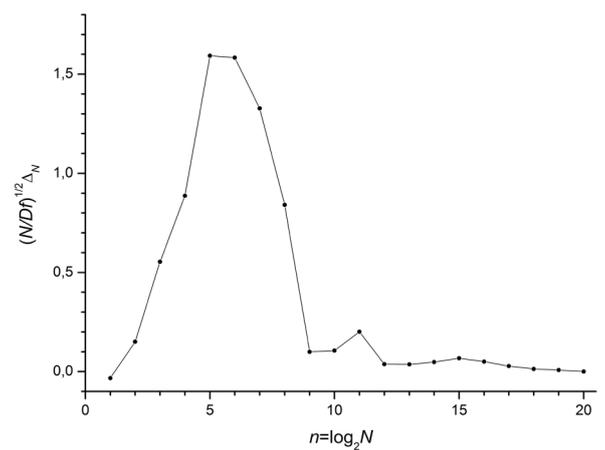


Рис.7. Зависимости нормированной погрешности $(N / Df)^{1/2} \Delta_N$ от $n = \log_2 N$ для последовательностей Соболя; $p = 7$.

Математическое ожидание погрешности равнялось бы нулю, то есть точки одинаково часто принимали бы положительные и отрицательные значения. В соответствии со свойствами распределения Гаусса 68% точек лежали бы в пределах ± 1 , 27% точек – в пределах ± 2 , а 5% точек выходили бы за эти пределы.

Однако для последовательностей Соболя мы наблюдаем существенно иную картину. Для $p = 2$ нормированная погрешность положительна и монотонно убывает, довольно быстро приближаясь к нулю при увеличении n .

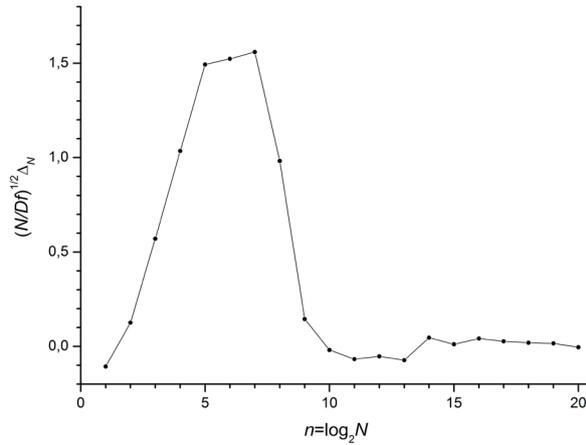


Рис.8. Зависимости нормированной погрешности $(N / Df)^{1/2} \Delta_N$ от $n = \log_2 N$ для последовательностей Соболя; $p = 8$.

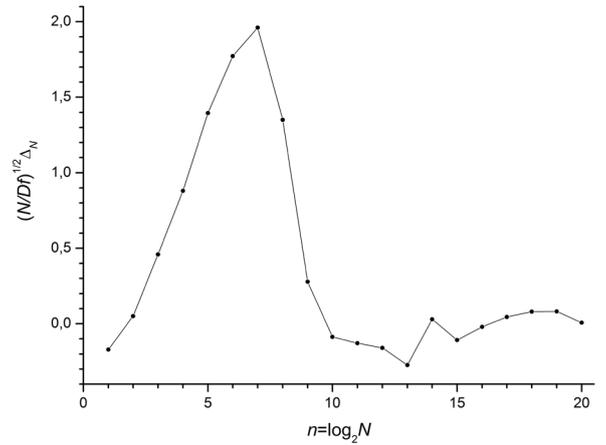


Рис.9. Зависимости нормированной погрешности $(N / Df)^{1/2} \Delta_N$ от $n = \log_2 N$ для последовательностей Соболя; $p = 9$.

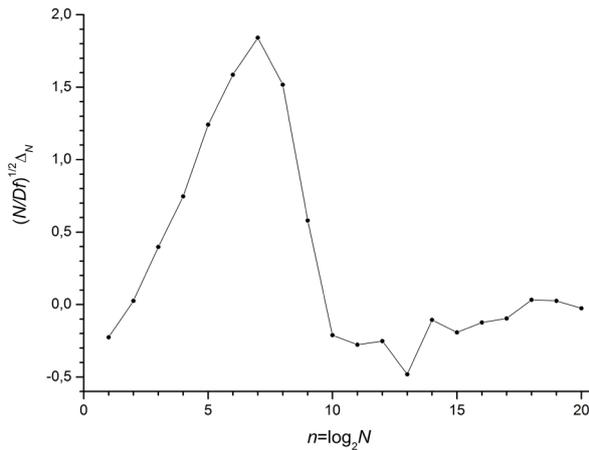


Рис.10. Зависимости нормированной погрешности $(N / Df)^{1/2} \Delta_N$ от $n = \log_2 N$ для последовательностей Соболя; $p = 10$.

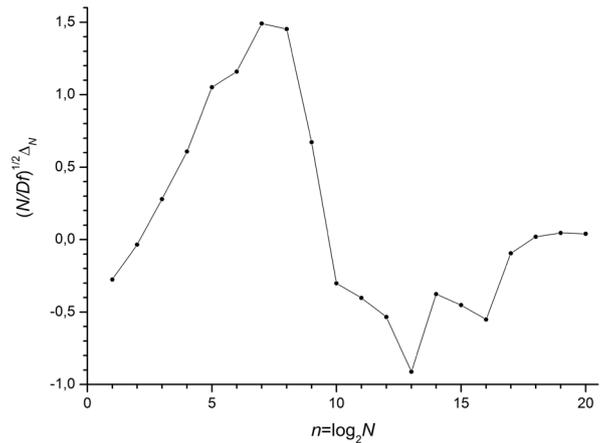


Рис.11. Зависимости нормированной погрешности $(N / Df)^{1/2} \Delta_N$ от $n = \log_2 N$ для последовательностей Соболя; $p = 11$.

Для $p = 3$ первые две точки выпадают, но начиная с $n = 3$ кривая также положительна и монотонно стремится к нулю; напомним, что $n = 3$ соответствует $N = 8$, что для расчетов методом Монте-Карло является чрезмерно малым (даже в грубых оценочных расчетах берут $N \geq 1000$). При $p = 4$ картина аналогичная, а монотонное падение начинается с $n = 4$ ($N = 16$). Для $p = 5$ убывание начинается с $n = 5$; при этом оно носит уже не вполне плавный характер. Для $p = 6$ качественное поведение еще несколько ухудшается: кривая после $n = 7$ не всюду убывает, а на небольшом ее участке

$n = 9 - 11$ наблюдается рост. При $p = 7$ качественное поведение еще хуже: появляется отрицательная погрешность при $n = 1$. При $p = 8$ уже в ряде точек погрешность отрицательна.

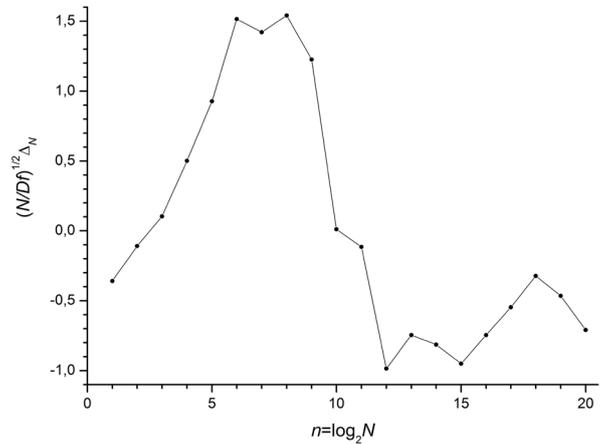
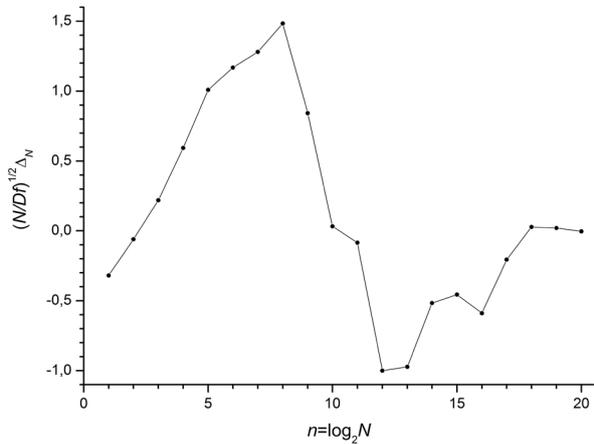


Рис.12. Зависимости нормированной погрешности $(N/Df)^{1/2} \Delta_N$ от $n = \log_2 N$ для последовательностей Соболя; $p = 12$.

Рис.13. Зависимости нормированной погрешности $(N/Df)^{1/2} \Delta_N$ от $n = \log_2 N$ для последовательностей Соболя; $p = 13$.

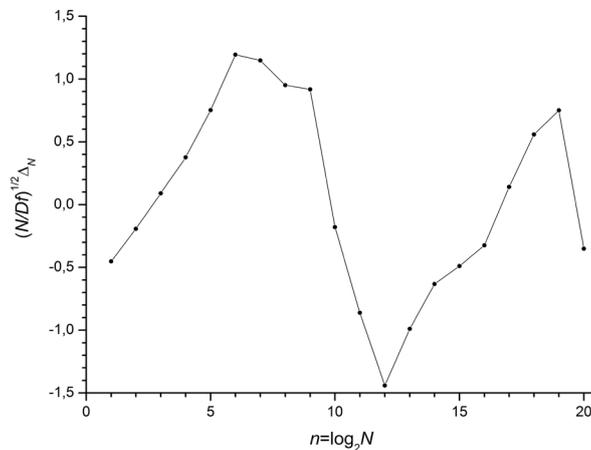


Рис.14. Зависимости нормированной погрешности $(N/Df)^{1/2} \Delta_N$ от $n = \log_2 N$ для последовательностей Соболя; $p = 16$.

Таким образом, при небольших p нормированная погрешность для последовательностей Соболя носит явно неслучайный характер и при $n > p$ монотонно убывает, стремясь к нулю. С увеличением размерности поведение погрешности становится все более хаотическим, хотя в целом стремление к нулю при больших n сохраняется.

2.3. Обработка графиков погрешности

Среднеквадратичная аппроксимация. Плавное убывание погрешности при небольших размерностях подсказывает возможность аппроксимации. Однако эта аппроксимация не может быть гладкой, типа интерполяционной аппроксимации Ричардсона для сеточных методов. В данном случае точки получаются статистическими методами, поэтому их обработку нужно вести с помощью среднеквадратичной аппроксимации. Для этого нужно выбрать вид аппроксимации и приписать точкам какие-то веса.

p	a	σ_a	b	σ_b	a / σ_a	$\Delta_{N=2^{20}}$
2	0.426E-5	3.251E-5	0.94	0.05	0.131	1.498E-6
3	0.111E-4	1.296E-4	1.15	0.19	0.086	2.246E-6
4	0.165E-4	2.738E-4	1.47	0.40	0.060	3.004E-6
5	0.278E-4	3.867E-4	1.54	0.56	0.072	3.882E-6
6	0.619E-4	6.200E-4	1.95	0.90	0.100	3.203E-6
7	0.938E-4	8.110E-4	2.07	1.17	0.116	2.955E-7
8	0.721E-4	9.730E-4	2.31	1.41	0.074	-9.362E-6
9	0.13E-3	1.28E-3	2.40	1.85	0.102	1.778E-5
10	-0.16E-3	1.42E-3	2.29	2.05	-0.113	-6.851E-5
11	-0.44E-3	1.50E-3	1.74	2.18	-0.293	1.154E-4
12	-0.71E-3	1.73E-3	1.71	2.51	-0.410	-1.087E-5
13	-3.52E-3	2.15E-3	1.74	3.12	-1.637	-2.62E-3
14	0.45E-3	2.41E-3	1.67	3.50	0.187	-1.58E-3
15	-0.07E-3	2.55E-3	1.19	3.70	-0.027	-2.07E-3
16	0.25E-3	2.85E-3	0.67	4.12	0.088	-1.81E-3

Таблица 1. Среднеквадратичная обработка погрешностей по формуле (7) с точками Соболя $n = 1 - 20$.

Из рис 2-6 видно, что при небольших p погрешность Δ_N убывает существенно быстрее, чем нормированная погрешность, поэтому в качестве рабочей гипотезы мы предположили закон убывания $\Delta_N \sim N^{-1}$. Но поскольку при увеличении p явно проявляется статистический характер погрешности, то стандартное уклонение этих погрешностей мы приняли пропорциональными в степени $N^{-1/2}$, поэтому мы опробовали следующий закон аппроксимации:

$$\Delta_N \approx a + bN^{-1}, \sigma_N \sim N^{-1/2}. \quad (7)$$

При каждом p проводились две обработки погрешностей методом наименьших квадратов. При первой обработке использовались все значения $n = 1 - 20$, а при второй – только достаточно большие значения, $n = 10 - 20$. При каждой обработке находились значения a, b и их стандартные уклонения. Эти значения представлены в табл 1 и 2. Подробно обсудим полученные результаты.

p	a	σ_a	b	σ_b	a / σ_a	$\Delta_{N=2^{20}}$
2	-0.636E-7	4.246E-7	1.520	0.014	-0.150	1.498E-6
3	-3.040E-7	2.991E-7	2.448	0.010	-1.016	2.246E-6
4	1.353E-6	2.773E-6	2.76	0.09	0.488	3.004E-6
5	3.514E-6	7.789E-6	3.42	0.26	0.451	3.882E-6
6	-1.199E-5	4.215E-5	9.98	1.38	-0.284	3.203E-6
7	1.549E-5	5.309E-5	9.08	1.74	0.292	2.955E-7
8	4.948E-5	5.485E-5	-3.48	1.80	0.902	-9.362E-6
9	1.426E-4	1.589E-4	-14.32	5.21	0.897	1.778E-5
10	-0.623E-4	2.672E-4	-33.41	8.76	-0.233	-6.851E-5
11	-1.730E-4	7.092E-4	-67.68	23.24	-0.244	1.154E-4
12	-0.493E-4	1.140E-3	-60.48	37.44	-0.432	-1.087E-5
13	-3.31E-3	1.32E-3	-66.23	43.36	-2.508	-2.62E-3
14	0.86E-3	1.55E-3	-107.21	50.86	0.555	-1.58E-3
15	0.39E-3	2.01E-3	-117.32	65.96	0.194	-2.07E-3
16	1.05E-3	2.26E-3	-181.05	73.97	0.465	-1.81E-3

Таблица 2. Среднеквадратичная обработка погрешностей по формуле (7) с точками Соболя $n = 10 - 20$.

Оценка погрешности. Рассмотрим столбцы a / σ_a в табл. 1 и 2. Видно, что все отношения, за исключением случая $p = 13$, лежат в пределах ± 1 , составляя в основном несколько десятых. Эти отношения проиллюстрированы на рис.15. Это показывает, что величина σ_a , несмотря на статистический характер погрешности, фактически является ее мажорантной оценкой. Поэтому именно величиной σ_a удобно пользоваться как оценкой точности результата. Выход отношения за эти пределы при $p = 13$ связан, скорее всего, с не очень удачными направляющими числами для этой размерности.

Из рис. 15 видно, что расчет в диапазоне $n = 10 - 20$ действительно выглядит статистическим, так как линия является резко изломанной с многократными переходами через нуль. Для диапазона $n = 1 - 20$ начало кривой выглядит плавным, что указывает на некоторую закономерность. Это является лишним аргументом в пользу того, чтобы не проводить усреднение по началу последовательности Соболя с $n < 10$.

Замечание. Выше мы говорили о величине a и ее стандарте. Они определялись для набора погрешностей Δ_N , однако погрешность $\Delta_N = I - I_N$. В практическом расчете нам неизвестно точное значение интеграла I , а тем самым и величина погрешности Δ_N . Однако если мы будем строить такую же

аппроксимацию формулой (7) не для величины Δ_N , а непосредственно для I_N , то величина a будет уточненным значением для I_N , а σ_a – стандартом этого уточнения. Приведенный выше анализ показывает, что величина σ_a является хорошей и, по-видимому, мажорантной оценкой для уточнения набора значений I_N методом наименьших квадратов с аппроксимацией (7). Поэтому в дальнейшем на графики оценки погрешностей мы будем наносить величину именно σ_a .

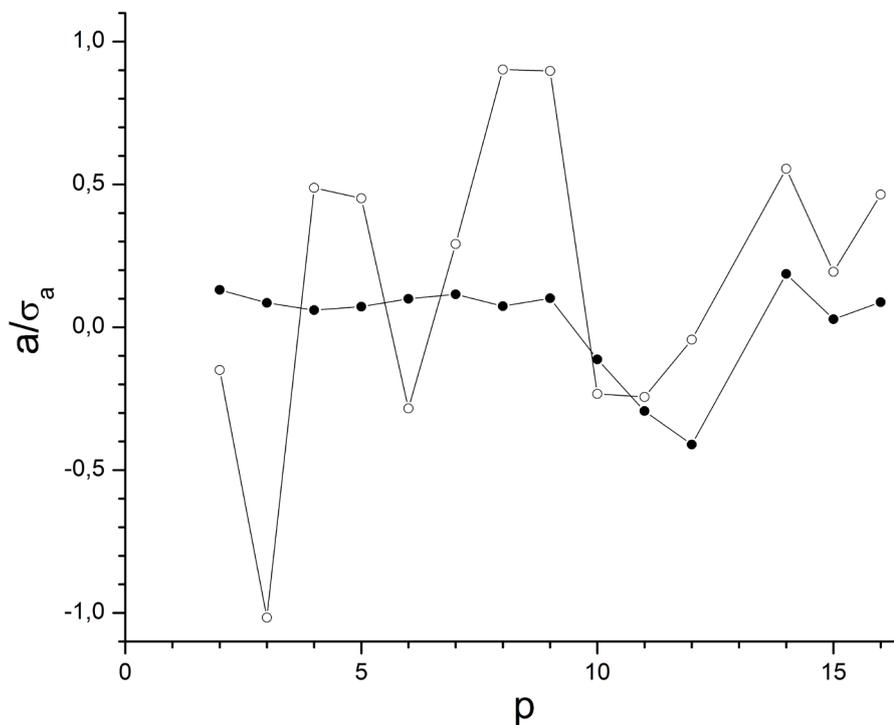


Рис.15. Зависимость отношения a к σ_a от p (не включено $p=13$); светлые кружки – набор $n = 10 - 20$, точки – набор $n = 1 - 20$.

Окончательный критерий. Поскольку при $N \rightarrow \infty$ аппроксимация (7) стремится к a , должно выполняться $Ma = 0$. Конкретная реализация a в среднеквадратичном расчете дает нам уточненную оценку погрешности. Если полученные значения a будут существенно меньше, чем погрешности на соответствующих рис. 2-14, то среднеквадратичная аппроксимация позволяет уточнить значение кубатуры.

Из табл.1 и 2 видно, что фактическая погрешность на последовательностях Соболя при $N = 2^{20}$ непривычно мала по сравнению с обычными результатами

методов Монте-Карло. Например, при $p=2$ она составляет $1.5 \cdot 10^{-6}$; с увеличением p она растет, но даже при $p=9$ составляет всего $2 \cdot 10^{-5}$. Само по себе это является отличным результатом. Однако мы смогли это установить, лишь имея в распоряжении тестовый интеграл с известным ответом. При вычислении произвольного интеграла мы не можем оценить погрешность.

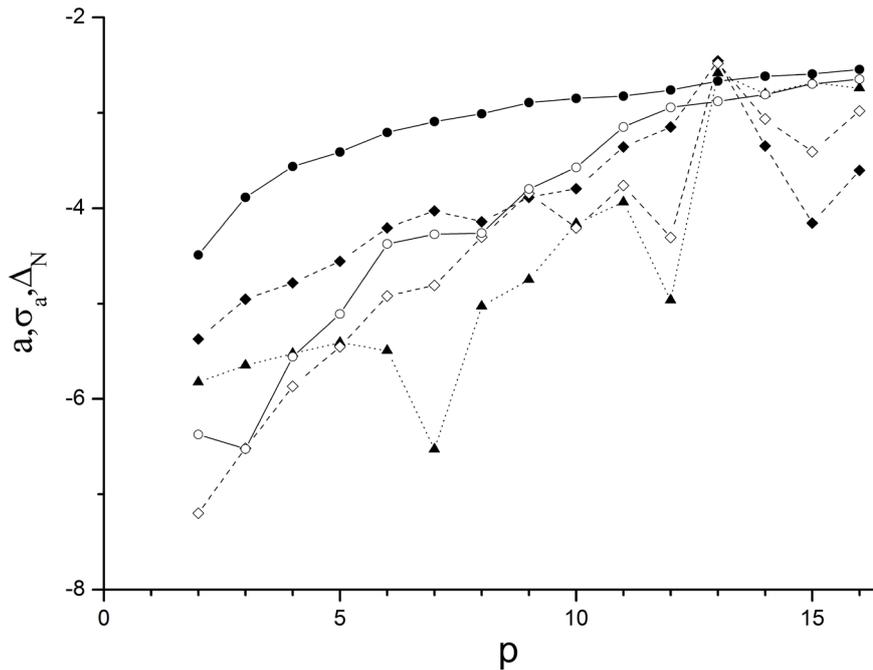


Рис.16. Зависимость погрешности от размерности p для $N = 2^{20}$: треугольник черный – $\lg |\Delta_N|$, черный ромб – $\lg |a|$ для диапазона $n = 1 - 20$ и черный кружок – $\lg \sigma_a$ для диапазона $n = 1 - 20$, светлый ромб – $\lg |a|$ для диапазона $n = 10 - 20$ и светлый кружок – $\lg \sigma_a$ для диапазона $n = 10 - 20$.

Величины a , даваемые среднеквадратичной обработкой интегралов с различными магическими N , также не могут являться оценкой погрешности: мы смогли их найти лишь сравнением с точным значением интеграла. Однако величины σ_a могут служить такой оценкой погрешности, так как они являются мажорантами для величин a . Поэтому проанализируем поведение σ_a ; значения этой величины приведены в табл. 1 и 2 и на рис. 16.

Если для среднеквадратичной обработки использован набор интегралов с во всеми магическими $N = 2^n$, $n = 1 - 20$, то соответствующая кривая σ_a лежит высоко. Если же при обработке откинута интегралы с малым числом точек

(оставлены $n = 10 - 20$), то кривая σ_a лежит существенно ниже. При малой размерности $p = 2$ разница составляет ~ 100 раз. При увеличении размерности разница сокращается, и при $p = 15$ кривые практически сливаются. Отсюда следует практический вывод – при проведении среднеквадратичного уточнения надо отбрасывать интегралы с $n < 10$.

Заметим, что зависимость обеих кривых σ_a от p плавная (почти гладкая), в отличие от кривой Δ_N . Это облегчает использование оценки по σ_a .

Влияние веса. В формуле (7) мы использовали два предположения. Первое касалось зависимости аппроксимации от N ; оно предполагало, что Δ_N убывает в среднем как N^{-1} . Оно основано на некоторых оценках, показывающих, что точки Соболя дают погрешность $O(N^{-1})$ вместо обычной статистической погрешности $O(N^{-1/2})$. Это предположение мы сохраним.

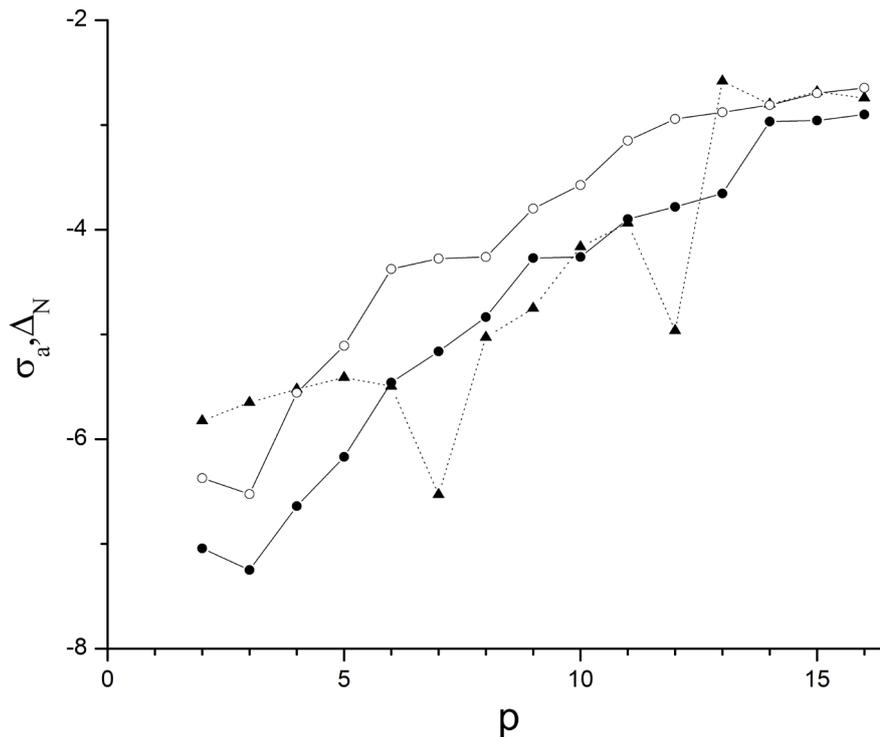


Рис.17. Расчет для $n = 10 - 20$; черный треугольник – $\lg |\Delta_N|$, черный кружок – $\lg \sigma_a$ для веса N^2 , светлый кружок – $\lg \sigma_a$ для веса N .

Однако второе предположение состоит в том, что вес для обработки соответствующей аппроксимации равен N ; это неявно соответствует

статистической погрешности $O(N^{-1/2})$. Кажется более логичным принять вес N^2 при статистической обработке величин Δ_N . Поэтому рассмотрим такой вариант обработки:

$$\Delta_N \approx a + bN^{-1}, \quad \sigma_N = N^{-1}. \quad (8)$$

Расчеты с такими весами также были проведены. При выбранном весе уже не столь важно, включаем ли мы в расчет начальные члены последовательности Δ_N с $N \leq 512$. Мы приведем результаты лишь для усеченного набора $N = 2^n$, $n = 10 - 20$. Они представлены на рис. 17.

На рис. 17 линии $\lg |\Delta_N|$ и $\lg \sigma_a$ для старого веса N перенесены с рис. 16. Добавлена линия $\lg \sigma_a$ для нового веса N^2 . Видно, что σ_a при новом весе нельзя считать хорошей оценкой погрешности: при $p < 6$ ее маркеры лежат существенно ниже, чем для фактической погрешности Δ_N . Поэтому вес N^2 для статистической обработки следует признать неудачным.

2.3. Смещенные магические точки

По-прежнему будет рассматривать только магические числа Соболя, так как кубатуры с немагическими числами дают существенно большую погрешность. Расположение точек Соболя слегка асимметрично. Например, если взять $N = 2^n$, то среднее арифметическое всех проекций точек на любую ось будет не 0.5, а $0.5(1 - 1/N)$. Очевидно, эта асимметричность неблагоприятна для получения хорошей точности кубатур.

На рис. 18 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 .

Мы провели расчеты для смещенных точек Соболя в том же диапазоне p и N , что и для несмещенных точек. К полученным интегралам применялась та же обработка методом наименьших квадратов, но при предложениях (8). Повышенный вес здесь представляется естественным, поскольку от смещенных точек Соболя следует ожидать более высокого порядка точности. Обсудим полученные результаты.

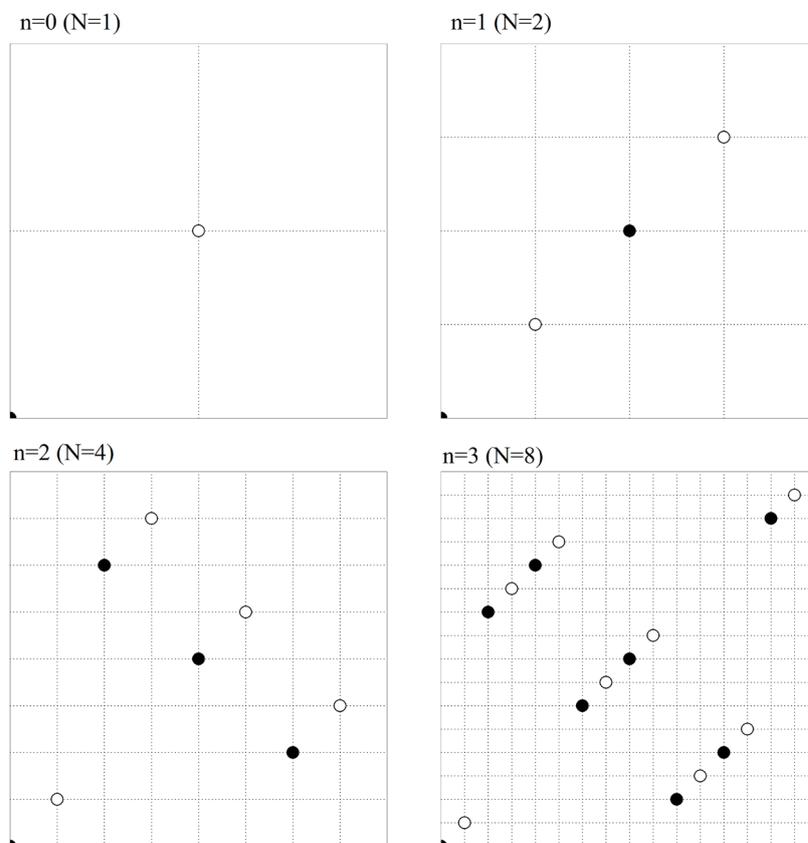


Рис.18. Магические точки Соболя для $p = 2$: точки – несмещенные, кружки – смещенные; значения n указаны около квадратов.

На рис. 19 маркером «светлый треугольник» показана погрешность для смещенных точек Соболя с $N = 2^{20}$, а маркером «светлый кружок» – σ_a по магическим числам $N = 2^{10} - 2^{20}$. Видно, что для больших размерностей p больше примерно шести эти величины отличаются несущественно (за

исключением $p = 12, 13$). При $p < 6$ σ_a систематически больше Δ_N , причем при $p = 2$ превышение составляет ~ 100 раз. При этом линия σ_a является плавной, а линия Δ_N – не плавной.

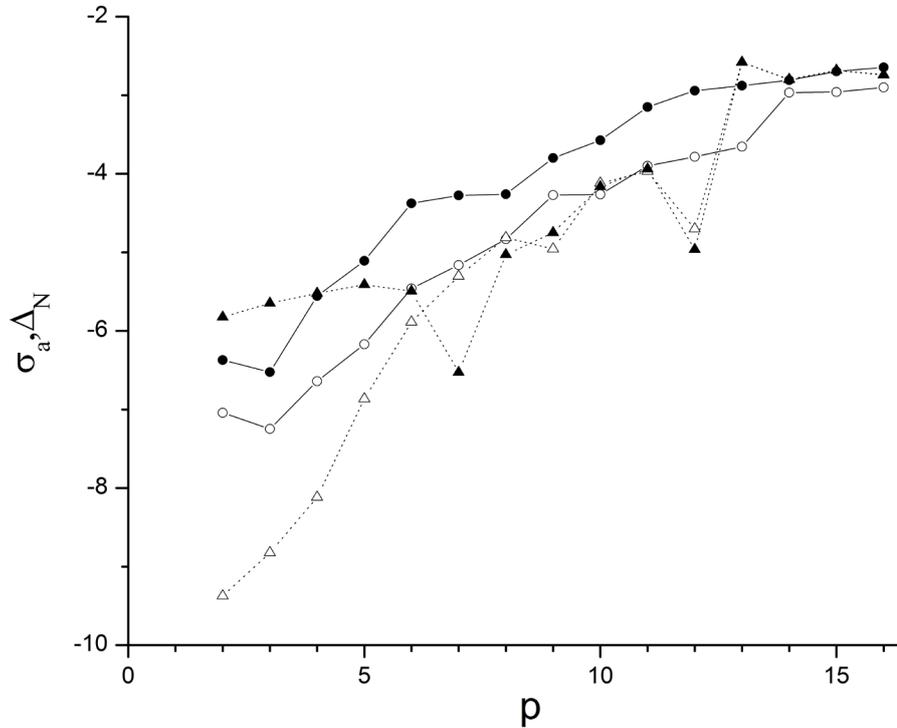


Рис.19. Логарифмы погрешностей и оценок: светлый треугольник – Δ_N и светлый кружок – σ_a для смещенных точек соболя, черный треугольник и черный кружок – то же для несмещенных точек.

Это позволяет дать следующую рекомендацию:

Для смещенных точек Соболя оценка погрешности по σ_a близка к фактической при p больше примерно шести, а при p меньше шести существенно превышает фактическую.

На рис. 19 черными маркерами показаны аналогичные погрешности для несмещенных точек Соболя. Видно, что погрешности Δ_N в обоих случаях близки при $p \geq 8$. При $p < 8$ смещенные точки Соболя дают существенный выигрыш в точности, причем при $p = 2$ он достигает $\sim 10^4$ раз! Это наглядно показывает преимущество смещенных точек Соболя. Мажорантную оценку σ_a также выгодней проводить по смещенным точкам Соболя.

Замечание. Погрешность Δ_N при $p = 13$ аномально велика. По-видимому, направляющие числа для этой размерности выбраны не вполне удачно. Однако величина σ_a для этой размерности не дает никакого выброса. Поэтому экстраполированной величиной a в этом случае можно пользоваться.

3. Сравнение различных методов

В разделе 1 описаны три качественно разных подхода к вычислению многомерных кубатур: регулярные кубатуры, классические методы Монте-Карло с псевдослучайными числами и последовательности Соболя. Из регулярных кубатур реально можно применять лишь произведение одномерных формул трапеций или средних; формулы средних предпочтительней по простоте записи и точности. Из псевдослучайных чисел в настоящее время наиболее популярным считается вихрь Мерсенна. Из последовательностей Соболя наилучшие результаты дают смещенные магические точки Соболя.

Сравним эти три подхода по величине погрешности при достаточно скромном числе точек $N = 2^{20}$, приблизительно равном 10^6 , на тесте (6). На рис. 20 приведены логарифмы погрешностей в зависимости от размерности p единичного куба. Проанализируем приведенные кривые.

3.1. Вихрь Мерсенна

В расчетах по классическим методам Монте-Карло математическое ожидание погрешности равно нулю, а стандарт определяется как $(Df / N)^{1/2}$. Подставляя сюда дисперсию теста (6), получим $\sigma_{МК} = 1 / N^{1/2} \left((\pi^2 / 8)^p - 1 \right)^{1/2}$, $N = 2^{20}$. Величина $\lg \sigma_{МК}$ плавно и слабо возрастает от -3.2 до -2.3 при увеличении p от 2 до 16. Следует ожидать, что расчеты для различных p будут хаотично колебаться вокруг этой кривой.

Расчеты были выполнены по стандартной программе. Запуски для различных p делались независимо, так что их результаты не должны были коррелировать друг с другом (если верить описанию стандартной программы). Однако соответствующие точки (черный перевернутый треугольник) на рис 20 образуют плавно осциллирующую кривую (наблюдается ~ 3 полуволны) с большой амплитудой ~ 1 . Это странный результат, заставляющий сомневаться

в правильности либо описания программы, либо самой программы. На последнее также указывает рис. 1 слева, где отчетливо видна сильная неоднородность распределения точек в двумерном случае.

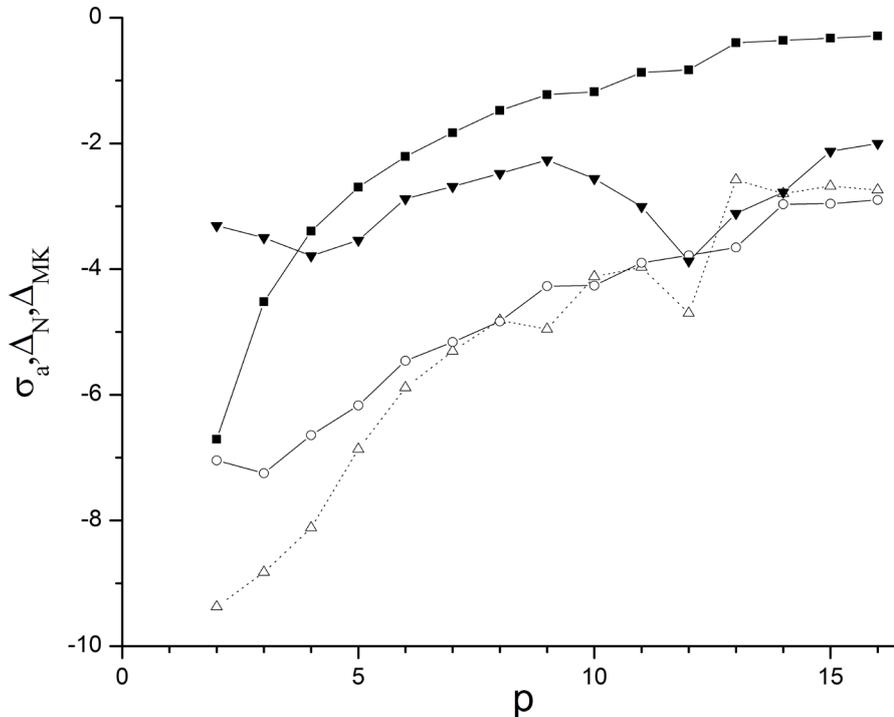


Рис.20. Логарифмы погрешностей для $N = 2^{20}$: светлый треугольник – Δ_N , кружок – σ_a для смещенных точек Соболя, черный перевернутый треугольник – вихрь Мерсенна, черный квадратик – метод средних.

3.2. Формула средних

Ее погрешность определяется формулой (2), где $k = N^{1/p}$ – число узлов по каждой координате. Тем самым ее погрешность есть $O(N^{-2/p})$; она должна быстро убывать с ростом p . Соответствующая кривая (маркер – черный квадрат на рис. 20) иллюстрирует высокую точность $\lg |R_{расч}| \approx 6.7$ 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 средних быстро становится неконкурентоспособным.

3.3. Последовательности Соболя

Лучшие результаты дают смещенные точки Соболя. Соответствующая кривая погрешности (маркер – светлый треугольник) лежит ниже всех линий на рис. 20. При $p = 2$ погрешность фантастически мала: $\lg |\Delta_N|$ приблизительно равен -9.4 . С увеличением p погрешность довольно быстро возрастает, но еще при $p = 6$ она очень мала по сравнению с другими методами: $\lg |\Delta_N|$ приблизительно равен -5.9 . Лишь при p приблизительно 13–16 погрешность становится близкой к классическим методам МК: $\lg |\Delta_N|$ приблизительно равен -2.7 . Таким образом, смещенные точки Соболя обеспечивают наилучшую точность при вычислении многомерных кубатур.

Важно отметить, что данное сравнение проводилось при весьма скромном числе точек N , приблизительно равном 10^6 . Обычно для расчетов многомерных кубатур берут гораздо большие числа точек. Но при увеличении N преимущество смещенных точек Соболя по сравнению с классическими методами МК усиливается. Поэтому можно утверждать, что смещенные точки Соболя кардинально превосходят все другие известные методы вычисления многомерных кубатур в единичном кубе.

Для полноты иллюстрации на рис. 20 приведен также стандарт σ_a для смещенных точек Соболя, найденный для условий (8). При $p \geq 6$ эта кривая хорошо совпадает с самими погрешностями, а при $p < 6$ является довольно сильно завышенной оценкой. Однако даже эта оценка существенно лучше, чем погрешность метода МК.

3.4. Рекомендация

Проведенное выше сравнение использовали для вычисления погрешности по точному решению задачи. В практических расчетах точное решение неизвестно. В этом случае мы получаем численные решения I_N и хотим узнать их погрешности. Как это сделать? Опишем соответствующую процедуру.

Возьмем последовательности Соболя с магическими числами $N = 2^n$. Для каждого N произведем смещение всех координат точек на величину $(2N)^{-1}$. Вычислим кубатуры по смещенным точкам и обозначим их через I_N . В качестве ответа возьмем I_N при максимальном использованном N . Для

мажорантной оценки его погрешности откинем интегралы для $N < 1024$ и совместно обработаем интегралы с N от 1024 до N_{\max} по формуле (8). Полученный при этом стандарт σ_a будет искомой мажорантной оценкой; эта оценка заметно завышена при $p < 6$, но будет почти оптимальна при $p \geq 6$.

Работа поддержана грантом РФФИ № 18-01-00175.

СПИСОК ЛИТЕРАТУРЫ

1. *И.М. Соболев* Численные методы Монте-Карло. М.: Наука, 1973.
2. *Д. Кнут*. Искусство программирования, том 2. Получисленные алгоритмы — М.: «Вильямс», 2007.
3. *А.А. Белов, Н.Н. Калиткин, М.А. Тинтул*. Визуальная верификация генераторов псевдослучайных чисел // Препринты ИПМ им. М. В. Келдыша, 2019, № 137.
4. *А.А. Белов, Н.Н. Калиткин, М.А. Тинтул*. Ненадежность известных генераторов псевдослучайных чисел // Журнал вычислительной математики и математической физики. 2020. Т. 60, № 11, С. 1807-1814.
5. *И.М. Соболев*. Равномерно распределенные последовательности с дополнительным свойством равномерности // Ж. вычисл. матем. и матем. физ., 1976, том 16, номер 5, 1332–1337
6. URL: <https://www.broda.co.uk/software.html>