Оптимальные дискретно-стохастические кубатурные формулы

Войтишек А. В.
Институт вычислительной математики и математической геофизики СО РАН, Новосибирск; vav@osmf.sscc.ru

1. Введение. В обзорном докладе [1] данной конференции описаны ситуации эффективного применения *дискретно-стохастических численных алгоритмов*, сочетающих в себе элементы сеточных схем и методов Монте-Карло. Упомянуты, в том числе, комбинированные алгоритмы численного интегрирования, в частности, дискретно-стохастические версии выборки по важности, выделения главной части, метода интегрирования по части области, метода сложной многомерной симметризации, метода равномерной выборки, метода с поправочным множителем, геометрического метода и др. (более подробно эти конструкции описаны в работах [2, 3]). В данном докладе проводится анализ алгоритмов из работы [4], которые можно трактовать как *дискретно-стохастическую версию выборки по группам.* Здесь удается получить оптимальные по порядку погрешности на классах гладких функций.

2. Алгоритмы Н. С. Бахвалова. Рассмотрим задачу интегрирования по единичному d-мерному кубу

$$I = \int_{Q_d} g(\mathbf{x})\,d\mathbf{x} = \int_0^1 \ldots \int_0^1 g(x^{(1)}, \ldots, x^{(d)})\,dx^{(1)} \ldots dx^{(d)}. \tag{1}$$

Разобьем исходную область интегрирования на равные кубы

$$X_m = \left\{ \mathbf{x} = \left(x^{(1)}, \ldots, x^{(d)}\right) : (j_m^{(i)} - 1)/\mu \le x^{(i)} \le j_m^{(i)}/\mu \right\}; \tag{2}$$

здесь $i = 1, \ldots, d;\ j_m^{(i)} = 1, \ldots, \mu;\ m = 1, \ldots, n;\ n = \mu^d$.

АЛГОРИТМ 1 [4]. *Реализуем по одной равномерно распределенной случайной точке* $\boldsymbol{\xi}_m$ *в каждом кубе* X_m *вида* (2) *и вычисляем приближение интеграла* (1) *согласно формуле*

$$I \approx \bar{\theta}_n = \frac{1}{\mu^d} \sum_{j_m^{(1)}, \ldots, j_m^{(d)} = 1}^{\mu} g(\boldsymbol{\xi}_m) = \frac{1}{n} \sum_{m=1}^{n} g(\boldsymbol{\xi}_m). \tag{3}$$

Рассмотрим также стандартный алгоритм метода Монте-Карло (см., например, [5]):

$$I \approx \bar{\zeta}_n = \frac{1}{n} \sum_{m=1}^{n} g(\hat{\boldsymbol{\xi}}_m), \tag{4}$$

где точки $\{\hat{\boldsymbol{\xi}}_m\}$ равномерно распределены во всем кубе Q_d. Несложно показать (см., например, [2, 5]), что для дисперсий оценок (3) и (4) выполнено неравенство $\mathbf{D}\bar{\theta}_n^{(M)} \le \mathbf{D}\bar{\zeta}_n$.

Рассмотрим класс $C^{\mathbf{r}}(\mathbf{L}; Q_d) = C^{\left(r^{(1)}, \ldots, r^{(d)}\right)}\left(L^{(1)}, \ldots, L^{(d)}; Q_d\right)$ функций d переменных, у которых $(r^{(s)} - 1)$-е производные по s-ой координате непрерывны, а $r^{(s)}$-е производные кусочно-непрерывны и ограничены по модулю константой $L^{(s)}$ в кубе Q_d для $s = 1, \ldots, d$. Предположим, что $g(\mathbf{x}) \in C^{(1,\ldots,1)}(L, \ldots, L; Q_d)$ (это означает, в частности, что $g(\mathbf{x})$ удовлетворяет условию Липшица с константой L по каждой из переменных).

Для выборочной дисперсии оценки (3) выполнено соотношение

$$\mathbf{D}\bar{\theta}_n = \sum_{m=1}^{n} \frac{\mathbf{D}(g(\boldsymbol{\xi}_m)}{n^2} = \sum_{m=1}^{n} \frac{\mathbf{E}(g(\boldsymbol{\xi}_m) - g_m)^2}{n^2}, \quad \text{где} \quad g_m = \mathbf{E}g(\boldsymbol{\xi}_m) = \mu^d \int_{X_m} g(\mathbf{x})\, d\mathbf{x}. \quad (5)$$

Из теоремы о среднем следует, что для каждого m найдется точка $\mathbf{x}_m \in X_m$ такая, что $\int_{X_m} g(\mathbf{x})\, d\mathbf{x} = g(\mathbf{x}_m)/\mu^d$. Следовательно, $g_m = g(\mathbf{x}_m)$. Для приращений $\Delta^{(i)}$, $i = 1, \ldots, d$ с учетом того, что $g(\mathbf{x}) \in C^{(1,\ldots,1)}(L, \ldots, L; Q_d)$, имеем

$$\left| g(x^{(1)} + \Delta^{(1)}, \ldots, x^{(d)} + \Delta^{(d)}) - g(x^{(1)}, \ldots, x^{(d)}) \right| \leq L(|\Delta^{(1)}| + \ldots + |\Delta^{(d)}|).$$

Так как точка $\boldsymbol{\xi}_m$ принадлежит X_m, то каждая из ее координат отличается от соответствующей координаты точки $\mathbf{x}_m$ не более чем на μ^{-1}. Поэтому из последнего неравенства, с учетом соотношения $\Delta^{(i)} < \mu^{-1}$, имеем

$$|g(\boldsymbol{\xi}_m) - g(\mathbf{x}_m)| \leq Ld\mu^{-1} \quad \text{или} \quad |g(\boldsymbol{\xi}_m) - g_m| \leq Ld\mu^{-1}. \quad (6)$$

Используя очевидное неравенство $\left| \int_Y z(\mathbf{y})\, d\mathbf{y} \right| \leq \sup_{\mathbf{y} \in Y} |z(\mathbf{y})| \times \operatorname{mes} Y$, из соотношений (5), (6) получаем

$$\mathbf{D}\bar{\theta}_n \leq \sum_{m=1}^{n} \frac{\sup_{\mathbf{y}_m \in X_m} (g(\mathbf{y}_m) - g_m)^2}{n^2} \leq \sum_{m=1}^{n} \frac{(Ld\mu^{-1})^2}{n^2} = n\, n^{-2}\, L^2\, d^2\, n^{-2/d} = \frac{L^2\, d^2}{n^{1+2/d}}.$$

Отсюда следует, что для погрешности $\delta_n^{(\theta)} = |I - \bar{\theta}_n|$ можно построить доверительный интервал вида

$$\mathbf{P}\left(\delta_n^{(\theta)} \leq H_\varepsilon \frac{\sigma^{(\theta)}}{\sqrt{n}} \leq H_\varepsilon \frac{Ld}{n^{1/2+1/d}} \right) \geq 1 - \varepsilon.$$

Здесь $H_\varepsilon = \text{const}$ и $\sigma^{(\theta)} = \sqrt{n\, \mathbf{D}\bar{\theta}_n}$, причем справедливо неравенство $\sigma^{(\theta)} \leq Ld/n^{1/d}$. Полученный порядок $t = 1/2 + 1/d$ является для погрешности $\tilde{\delta}_n^{(M)} \sim n^{-t}$ неулучшаемым. Это дает следующее утверждение. Обозначим через Ξ_n произвольную кубатурную формулу, использующую значения подынтегральной функции $g(\mathbf{x})$ в n узлах $\{\tilde{\boldsymbol{\xi}}_1, \ldots, \tilde{\boldsymbol{\xi}}_n\}$ – случайных или фиксированных (приближения (3) и (4) являются частными примерами формулы Ξ_n).

УТВЕРЖДЕНИЕ 1 [4]. *Существуют положительные константы $H(\mathbf{r}, \mathbf{L})$ и P, удовлетворяющие следующему соотношению. Для любого натурального n и любой формулы Ξ_n найдется функция $g(\mathbf{x}) \in C^{\mathbf{r}}(\mathbf{L}; Q_d)$, для которой*

$$\delta_n^{(\Xi)}(\mathbf{r}, \mathbf{L}) = |\Xi_n - I| > \frac{H(\mathbf{r}, \mathbf{L})}{n^{r+1/2}} \quad (7)$$

с вероятностью P; здесь $1/r = 1/r^{(1)} + \ldots + 1/r^{(d)}$.

Для рассматриваемого нами множества функций имеем $\mathbf{r} = (1, \ldots, 1)$, $\mathbf{L} = (L, \ldots, L)$ и $r = 1/d$. Следовательно, согласно соотношению (7) для класса $C^{(1,\ldots,1)}(L, \ldots, L; Q_d)$ получаем $\delta_n^{(\theta)} > \tilde{H}/n^{1/2+1/d}$. Таким образом, для оценки (3) нашлись константы $H_1(P)$ и $H_2(P)$, для которых с заданной вероятностью P выполнено двойное неравенство $H_1(P)/n^{1/2+1/d} < \delta_n^{(\theta)} \leq H_2(P)/n^{1/2+1/d}$, что означает оптимальность алгоритма 1 по порядку t вероятностной погрешности $\delta_n^{(\theta)} \sim n^{-t}$.

Аналогичным образом можно показать, что если подынтегральная функция $g(\mathbf{x})$ принадлежит классу $C^2(L, Q_d) \subset C^{(2,..,2)}(L, .., L; Q_d)$ функций, имеющих непрерывные смешанные производные второго порядка, ограниченные в совокупности общей константой L, то оптимальный порядок $t = 1/2{+}2/d$ дает следующая кубатурная формула.

АЛГОРИТМ 2 [4]. *Реализуем по одной случайной точке* $\boldsymbol{\xi}_m$ *в каждом кубе* X_m *вида* (2). *Строим точку* $\boldsymbol{\xi}_{m,sim}$, *симметричную* $\boldsymbol{\xi}_m$ *относительно центра куба* X_m. *Вычисляем приближение интеграла* (1) *согласно формуле*

$$I \approx \bar{\Theta}_n = \frac{1}{n} \sum_{m=1}^{n/2} \left(g(\boldsymbol{\xi}_m) + g(\boldsymbol{\xi}_{m,sim}) \right). \tag{8}$$

Заметим, что алгоритмы 1 и 2 допускают несколько названий. Во-первых, это частные случаи *алгоритма расслоенной выборки* (см., например, [5]), для которых число испытаний n либо равно числу M подмножеств (2), либо $n = 2M$ соответственно. Во-вторых, в связи с наличием «сеточного» разбиения области интегрирования X на малые кубы X_m вида (2) и выбора одной или двух случайных точек в каждом малом кубе можно отнести алгоритмы 1 и 2 к *дискретно-стохастическим алгоритмам численного интегрирования* [2]. В-третьих, алгоритмы 1 и 2 являются частными случаями *случайных кубатурных формул* (см., например, [5]). Отметим, что представленные здесь результаты из работы [4] вызвали большой научный резонанс и привели к бурному развитию (главным образом, в западноевропейских научных школах) *теории сложности численных алгоритмов* [6]. Классической задачей в этой теории является следующая: *при фиксированном числе операций n определить максимальный порядок t погрешности $\delta_n \sim n^{-t}$ (в обычном или вероятностном смысле) заданного класса вычислительных алгоритмов*. Алгоритмы численного интегрирования являются наиболее удачными иллюстрациями конструкций и методик теории сложности.

3. О трудоемкости алгоритмов Н. С. Бахвалова. На самом деле, максимальные порядки t по n погрешностей δ_n рассмотренных выше алгоритмов 1 и 2 не гарантируют того, что кубатурные формулы (2) и (8) действительно являются *оптимальными* (т. е. дающими заданный уровень погрешности за минимальное время) из-за необходимости получения случайных точек $\boldsymbol{\xi}_m$ в подмножествах X_m вида (2); это требует осуществления многократных обращений к датчику стандартных случайных чисел α_i, равномерно распределенных в интервале $(0, 1)$. Как известно (см., например. [5]), при использовании широко распространенных датчиков, основанных на методе вычетов, получение α_i является достаточно «дорогостоящей» компьютерной операцией. Описанные затруднения проявляются более заметно с ростом кратности d интеграла (1).

Нами было проведено численное сравнительное исследование алгоритмов 1 и 2 с точки зрения их трудоемкости при получении заданного уровня погрешности. Формулы (3) и (8) сравнивались как с «детерминированными» численными схемами (типа формулы прямоугольников, для которой значения подынтегральной функции вычислялись в центральных точках кубов X_m), так и с простейшим методом Монте-Карло (4). Рассматривалась также специальная модификация алгоритма (4)

– *дискретно-стохастическая версия выборки по важности* [2, 3]. При проведении численных экспериментов использовалась *стохастическая тестовая система* [7], основанная на использовании модельных траекторий случайных полей.

На основе численных экспериментов (частично представленных в работе [3]) можно сформулировать следующие выводы.

1. Для размерностей $d < 5$ вычисление интегралов с помощью оптимальных алгоритмов 1 *и* 2 *от всех тестовых функций более экономично.*

2. При вычислении интегралов от достаточно «гладких» функций из стохастической системы функций и размерностей $d \leq 7$ время работы дискретно-стохастической версии выборки по важности сравнимо с затратами алгоритмов 1 *и* 2*. Но для больших размерностей интеграла и функций малой гладкости этот метод малоэффективен.*

3. При вычислении интегралов высокой размерности ($d \geq 6$) трудоемкость простейшего метода Монте-Карло (4) *не превосходит трудоемкости «оптимальных» алгоритмов* 1 *и* 2*.*

4. Трудоемкость алгоритмов 1 *и* 2 *во всех случаях не превосходит трудоемкости метода прямоугольников. При небольших размерностях интеграла метод прямоугольников эффективнее простейшего метода Монте-Карло* (4).

Таким образом, смешанные дискретно-стохастические численные схемы (алгоритм выборки по важности, алгоритмы 1 и 2) могут быть эффективными для «промежуточных» размерностей ($d = 3 - 9$). Для малых размерностей и гладких подынтегральных функций неплохо работают простейшие кубатурные формулы (вплоть до формулы прямоугольников). В ходе численных экспериментов нами было также установлено, что методы Бахвалова (алгоритмы 1 и 2), предложенные им для пространств функций $C^{(1,..,1)}(L, \dots, L; Q_d)$ и $C^2(L; Q_d)$, можно применять и для кусочно-непрерывных функций с кусочно-непрерывными производными.

Работа выполнена при финансовой поддержке РФФИ (проекты 09–01–00035, 10–01–00040, 11–01–12030).

ЛИТЕРАТУРА

1. *Войтишек А. В.* Теория и приложения дискретно-стохастических численных методов: новые результаты // Данный сборник. С....

2. *Войтишек А. В.* Дискретно-стохастические модификации стандартного метода Монте-Карло. Новосибирск: НГУ, 2009.

3. *Каблукова Е. Г.* Адаптивные дискретно-стохастические алгоритмы численного интегрирования // Дисс. на соиск. уч. степени канд. физ.-матем. наук. Новосибирск: ИВМиМГ СО РАН, 2009.

4. *Бахвалов Н. С.* Об оптимальных оценках скорости сходимости квадратурных процессов и методов интегрирования типа Монте-Карло на классах функций // Численные методы решения дифференциальных и интегральных уравнений и квадратурные формулы. М.: Наука, 1964. С. 5–63.

5. *Михайлов Г. А., Войтишек А. В.* Численное статистическое моделирование. Методы Монте-Карло. М.: Изд. центр «Академия», 2006.

6. *Traub J. F., Wasilkowski G. W. and Wozniakowski H.* Information-based Complexity. New York: Academic Press, 1988.

7. *Войтишек А. В., Каблукова Е. Г., Булгакова Т. Е.* Использование спектральных моделей случайных полей при исследовании алгоритмов численного интегрирования // Вычислительные технологии. 2004. Т. 9, специальный выпуск. С. 50–61.