

О выборе плотностей распределения узлов адаптивных сеток в стохастических алгоритмах многомерного численного интегрирования и приближения функций

А. В. Войтишек*, Д. А. ПРАСОЛ

Институт вычислительной математики и математической геофизики СО РАН,
Новосибирск, Россия

*Контактный e-mail: vav@osmf.ssc.ru

Сформулирован аналог принципа эквираспределения для стохастических адаптивных сеток, используемых при решении многомерных задач численного интегрирования. Рассмотрена проблема выбора плотности распределения узлов при реализации алгоритма самоорганизации Т. Кохонена для построения адаптивных сеток при решении задачи численного приближения функций. Для гладких приближаемых функций в одномерном случае предложено использовать в качестве плотности адаптивных узлов корень квадратный из модуля второй производной приближаемой функции. Приведен пример функции, для которой данная рекомендация приводит к улучшению аппроксимационных свойств сеточного приближения.

Ключевые слова: адаптивная сетка, принцип эквираспределения, вероятностная плотность распределения, метод выборки по важности, схема Т. Кохонена, выбор плотности узлов при приближении функций.

Введение

При численном решении ряда прикладных задач целесообразно использовать адаптивные сетки, в которых сгущения узлов в подобластях позволяют более эффективно получать приближенные решения.

В данной работе обсуждаются применение и численная реализация адаптивных сеток в алгоритмах численного интегрирования и аппроксимации функций. Особое внимание уделено решению многомерных задач, в которых целесообразно применять стохастические (вероятностные) алгоритмы построения адаптивных сеток.

В разд. 1 приведены рассуждения о том, что в основном принципе построения адаптивных сеток — принципе эквираспределения [1] — так называемая управляющая функция имеет основные свойства вероятностной плотности распределения [2]. Сформулирован вероятностный аналог принципа эквираспределения (в интегральной форме) для случая, когда управляющая функция в точности является вероятностной плотностью, индуцирующей неупорядоченный адаптивный набор точек (как в алгоритме выборки по важности [3]).

В разд. 2 приведен краткий анализ использования стохастических адаптивных сеток при построении квадратурных и кубатурных формул для приближенного вычисления интегралов [3, 4], в том числе рассмотрены особенности компьютерной реализации упомянутого выше алгоритма выборки по важности.

В разд. 3 приведены соображения об использовании стохастического алгоритма построения адаптивных сеток — схемы Т. Кохонена [5, 6] — в задаче приближения функций, а также сформулированы рекомендации по выбору плотности распределения “узлов притяжения” в этом алгоритме.

Отметим особенность представления материала в данной работе. Хотя речь идет об алгоритмах, особенно эффективных для многомерных задач, ряд принципиальных аналитических рассуждений и рекомендаций в тексте сделан для одномерного случая. Перенос соответствующих суждений на многомерный случай технически не сложен (соответствующие примеры приведены в заключении данной статьи), однако он требует использования “тяжеловесных” многомерных обозначений и рассуждений, что приводит к затруднениям при восприятии основных идей данной статьи.

1. Вероятностный аналог принципа эквираспределения

При построении адаптивных сеток, как правило, реализуется тот или иной вариант принципа эквираспределения [1]. Этот принцип связан с построением невырожденного отображения $\mathbf{x} = \mathbf{x}(\mathbf{y})$, которое переводит зафиксированную (чаще всего равномерную) сетку на “вычислительной” области Q в адаптивную сетку на заданной “физической” области X , и удовлетворяет соотношению

$$f(\mathbf{x}(\mathbf{y}))J(\mathbf{y}) \equiv H_1 = \text{const}, \quad (1.1)$$

где $J(\mathbf{y})$ — якобиан отображения; H_1 — положительная константа; $f(\mathbf{x})$ — специально выбранная *управляющая функция*, такая, что

$$f(\mathbf{x}) \geq 0 \quad \text{при} \quad \mathbf{x} \in X \quad \text{и} \quad f(\mathbf{x}) = 0 \quad \text{при} \quad \mathbf{x} \notin X. \quad (1.2)$$

На практике обычно рассматриваются случаи, когда $X \subseteq \mathbf{R}^d$, а $Q \subseteq \mathbf{R}^s$, где $1 \leq s \leq d$.

В дальнейшем будем полагать, что функция $f(\mathbf{x})$ непрерывна или кусочно-непрерывна, и выполнено соотношение

$$\int f(\mathbf{x}) d\mathbf{x} = 1. \quad (1.3)$$

Свойствами (1.2), (1.3) обладает, в частности, вероятностная плотность распределения $f_{\boldsymbol{\xi}}(\mathbf{x})$ некоторого случайного вектора (d -мерной случайной величины) $\boldsymbol{\xi}$ [2]. Но более существенным свойством вероятностной плотности является то, что она определяет “сгущения” выборочных значений (реализаций) $\boldsymbol{\xi}_j$ случайного вектора $\boldsymbol{\xi}$.

Управляющая функция $f(\mathbf{x})$, обладающая свойствами (1.1) — (1.3), также определяет “сгущения” узлов реализуемой с ее помощью адаптивной сетки.

Покажем это для простейшего одномерного случая, когда $d = 1$ и

$$\mathbf{x} = x \in X \subset \mathbf{R} \quad \text{и} \quad \mathbf{y} = y \in Q = [0, 1] \subset \mathbf{R}.$$

Построение адаптивной сетки

$$\tilde{X}^{(M)} = \{x_0 < x_1 < x_2 < \dots < x_{M-1} < x_M\}, \quad (1.4)$$

где

$$x_j = x(y_j), \quad y_j = j\Delta y, \quad \Delta y = 1/M, \quad j = 0, 1, \dots, M, \quad (1.5)$$

в этом случае связано с рассмотрением частного случая соотношения (1.1)

$$f(x(y))x'_y(y) = H_2 = \text{const}. \quad (1.6)$$

Здесь $x'_y(y)$ обозначает соответствующую производную по переменной y (это в данном случае якобиан отображения $x = x(y)$).

Интегрируя это соотношение на отрезке $[y_j, y_{j+1}]$ и предполагая непрерывность функции $f(x)$ на этом отрезке, получаем

$$\int_{x_j}^{x_{j+1}} f(z) dz = H_2(y_{j+1} - y_j) \quad (1.7)$$

или

$$f(\tilde{x}_j)(x_{j+1} - x_j) = H_3, \quad (1.8)$$

где $\tilde{x}_j \in [x_j, x_{j+1}]$, $H_3 = H_2\Delta y = \frac{H_2}{M}$, $j = 0, 1, \dots, M - 1$.

Таким образом, при заданном количестве узлов M для подобластей, в которых функция $f(x)$ велика, длины отрезков $[x_j, x_{j+1}]$ малы.

Учитывая соотношения (1.2), (1.3), (1.8), функцию $f(x)$ можно назвать *плотностью распределения узлов адаптивной сетки*.

Теперь рассмотрим одномерную неупорядоченную стохастическую адаптивную сетку для метода Монте-Карло (см. [3], а также разд. 2 данной работы)

$$\Xi^{(n)} = \{\xi_1, \dots, \xi_n\},$$

где ξ_i — независимые выборочные значения случайной величины ξ , имеющей плотность распределения $f_\xi(z)$, а также соответствующий аналог левой части соотношения (1.7) (которое можно назвать интегральной формой принципа эквираспределения):

$$I^{(i)} = \int_{\xi_i}^{\xi_{i+1}} f_\xi(z) dz = F_\xi(\xi_{i+1}) - F_\xi(\xi_i), \quad i = 1, \dots, n - 1.$$

Здесь $F_\xi(x) = \mathbf{P}\{\xi < x\} = \int_{-\infty}^x f_\xi(z) dz$ — функция распределения случайной величины ξ [2].

Хорошо известно [3], что случайная величина

$$\alpha = F_\xi(\xi)$$

имеет равномерное распределение в интервале $(0, 1)$.

Таким образом, случайная величина $I^{(i)}$ равна

$$\int_{\xi_i}^{\xi_{i+1}} f_\xi(z) dz = \alpha^{(1)} - \alpha^{(2)} \quad (1.9)$$

для любой плотности распределения $f_{\xi}(z)$ и любых случайных узлов ξ_i и ξ_{i+1} . Здесь $\alpha^{(1)}$ и $\alpha^{(2)}$ — независимые стандартные случайные величины [3], равномерно распределенные в интервале $(0, 1)$.

Используем следующее утверждение.

Лемма [2]. Сумма $\eta^{(1)} + \eta^{(2)}$ двух независимых случайных величин $\eta^{(1)}$ и $\eta^{(2)}$ с плотностями распределения $f^{(1)}(z)$ и $f^{(2)}(z)$ имеет плотность распределения в виде свертки

$$f_{\eta^{(1)}+\eta^{(2)}}(z) = (f_{\eta^{(1)}} * f_{\eta^{(2)}})(z) = \int f^{(1)}(u)f^{(2)}(z-u) du.$$

Из леммы следует, что случайная величина $I^{(i)}$ имеет “универсальное” (для различных $i = 1, \dots, n-1$) распределение с плотностью

$$f_{I^{(i)}}(z) = f_{\alpha^{(1)}+(-\alpha^{(2)})}(z) = \beta^{(1)}(z) = \begin{cases} 1+z & \text{для } -1 \leq z \leq 0, \\ 1-z & \text{для } 0 \leq z \leq 1, \\ 0 & \text{иначе.} \end{cases} \quad (1.10)$$

Здесь $\beta^{(1)}(z)$ — B -сплайн первого порядка (или “функция-крышка” [7]).

Таким образом, соотношение (1.9) можно назвать *вероятностным аналогом принципа эквивалентности*.

2. Использование адаптивных сеток в стохастических алгоритмах численного интегрирования

Рассмотрим задачу приближенного вычисления многократного интеграла

$$I = \int_X g(\mathbf{x}) d\mathbf{x} = \int_{\mathbf{R}^d} g(\mathbf{x}) d\mathbf{x} = \int g(\mathbf{x}) d\mathbf{x}.$$

Известно [3, 4], что для достаточно больших размерностей d (конкретнее, для $d \geq 4$ [3]) наиболее эффективным алгоритмом вычисления интеграла I является весовой метод Монте-Карло, основанный на представлении интеграла в виде математического ожидания

$$I = \int \frac{g(\mathbf{x})}{f_{\xi}(\mathbf{x})} f_{\xi}(\mathbf{x}) d\mathbf{x} = \mathbf{E}\zeta, \quad \zeta = \frac{g(\xi)}{f_{\xi}(\xi)},$$

где случайный вектор ξ распределен согласно выбираемой плотности $f_{\xi}(\mathbf{x})$ с последующим численным приближением интеграла I в соответствии с законом больших чисел [2, 3]

$$I \approx \frac{1}{n} \sum_{i=1}^n \zeta_i, \quad \zeta_i = \frac{g(\xi_i)}{f_{\xi}(\xi_i)}. \quad (2.1)$$

Здесь ξ_1, \dots, ξ_n — реализуемые на компьютере выборочные значения случайного вектора ξ .

Одно из принципиальных преимуществ метода (2.1) состоит в том, что он допускает корректную аналитическую оптимизацию на основе принципа “тот алгоритм эффективнее, который позволяет получить заданный уровень погрешности за меньшее

время счета”, в то время как “детерминированные” методы численного интегрирования — квадратурные и кубатурные формулы — сравниваются (не вполне корректно) только по зависимости погрешности от шага сетки [4].

Затраты (время счета) алгоритма (2.1) равны величине

$$s = nt, \quad (2.2)$$

где t — среднее время для получения выборочного значения $\zeta_i = g(\xi_i)/f_{\xi}(\xi_i)$. Как несложно показать [3], при фиксированном уровне погрешности величина (2.2) пропорциональна величине трудоемкости

$$S = \mathbf{D}\zeta \cdot t, \quad (2.3)$$

где $\mathbf{D}\zeta$ — дисперсия случайной величины (“весовой оценки”) ζ .

В свою очередь, при оптимизации алгоритма (2.1) на основании минимизации величины (2.3) широко используется *принцип выборки по важности*, состоящий в том, что среди алгоритмов (2.1), имеющих примерно одинаковое среднее время t реализации выборочных значений ζ_i , тот имеет меньшую трудоемкость S из (2.3), для которого

$$f_{\xi}(\mathbf{x}) \approx H_4 |g(\mathbf{x})|, \quad H_4 = \frac{1}{\int |g(\mathbf{u})| d\mathbf{u}}. \quad (2.4)$$

Термин “выборка по важности” соответствует английскому термину *importance sampling*. Такое название объясняется тем, что если $f_{\xi}(\mathbf{x})$ пропорциональна модулю подынтегральной функции $g(\mathbf{x})$, то в частях области X , где $|g(\mathbf{x})|$ больше и вклад которых в интеграл I более существен, будет выбираться больше случайных точек $\{\xi_i\}$.

Таким образом, в алгоритме выборки по важности (2.1), (2.4) реализуется *адаптивная стохастическая сетка*

$$\Xi^{(n)} = \{\xi_1, \dots, \xi_n\}. \quad (2.5)$$

Что касается конкретных алгоритмов компьютерной реализации выборки (2.5), то здесь можно воспользоваться соображениями из разд. 3.2.4 учебника [3] (более детально они представлены в [8]). Эти соображения состоят в том, что в качестве плотности $f_{\xi}(\mathbf{x})$ целесообразно использовать приближения модуля подынтегральной функции $g(\mathbf{x})$ вида

$$f_{\xi}(\mathbf{x}) = H_5 L_M |g(\mathbf{x})| = H_5 \sum_{j=1}^M w_j(\mathbf{g}) \chi_j(\mathbf{x}), \quad (2.6)$$

где $\{\chi_j(\mathbf{x})\}$ — заданные базисные функции, согласованные с заданными узловыми точками $\{\mathbf{x}_1, \dots, \mathbf{x}_M\}$, коэффициенты $\{w_j(\mathbf{g})\}$ зависят от значений $\mathbf{g} = (|g(\mathbf{x}_1)|, \dots, |g(\mathbf{x}_M)|)$, а H_5 — нормирующая константа.

Здесь возникают требования моделируемости приближения $L_M |g(\mathbf{x})|$ [3, 8]. Из соотношения (2.6) следует, что для численного моделирования адаптивных случайных узлов (2.5) можно использовать алгоритм метода суперпозиции [3].

Алгоритм 1.

1. Используя один из алгоритмов моделирования целочисленной случайной величины η (см., например, [3]) с распределением

$$\mathbf{P}\{\eta = j\} = p_j, \quad p_j = \frac{H_5 w_j(\mathbf{g})}{H_6^{(j)}}, \quad H_6^{(j)} = \int \chi_j(\mathbf{u}) d\mathbf{u}, \quad j = 1, \dots, M,$$

реализуем выборочное значение $\eta_i = t$, $i = 1, \dots, n$.

2. Реализуем выборочное значение ξ_i согласно плотности

$$f_m(\mathbf{x}) = H_6^{(m)} \chi_m(\mathbf{x}), \quad \mathbf{x} = (x^{(1)}, \dots, x^{(d)}) \in \mathbf{R}^d. \quad (2.7)$$

Для корректной реализации алгоритма 1 требуется неотрицательность коэффициентов $\{w_j(\mathbf{g})\}$. Кроме того, базисные функции $\{\chi_j(\mathbf{x})\}$ должны быть пропорциональными плотностям распределения случайных векторов вида (2.7), для которых имеются эффективные алгоритмы реализации выборочных значений. Последние два требования определяют моделируемость приближения $L_M|g(\mathbf{x})|$ из соотношения (2.6).

Проведенные в работе [8] исследования известных приближений вида (2.6) показали, что требованиям моделируемости наилучшим образом удовлетворяет конечно-элементная аппроксимация Стренга — Фикса [7]. Базисные функции строятся следующим образом.

Предполагается, что в \mathbf{R}^d задана равномерная прямоугольная сетка с шагом h и каждый узел \mathbf{x}_j , используемый в аппроксимации (2.6), можно описать как

$$\mathbf{x}_j = \left(k_{(j)}^{(1)} h, \dots, k_{(j)}^{(d)} h \right),$$

где $k_{(j)}^{(s)}$ — целые числа. Соответствующая узлу \mathbf{x}_j базисная функция $\chi_j(\mathbf{x})$ имеет специальный вид

$$\chi_j(\mathbf{x}) = \beta^{(r)} \left(\frac{x^{(1)}}{h} - k_{(j)}^{(1)} \right) \cdot \dots \cdot \beta^{(r)} \left(\frac{x^{(d)}}{h} - k_{(j)}^{(d)} \right). \quad (2.8)$$

Здесь $\beta^{(r)}(z)$ — B -сплайн r -го порядка.

Для базиса Стренга — Фикса (2.8) плотность (2.7), используемая в алгоритме 1, имеет вид

$$f_m(\mathbf{x}) = f_m^{(1)}(x^{(1)}) \cdot \dots \cdot f_m^{(d)}(x^{(d)}), \quad f_m^{(s)}(x^{(s)}) = \frac{1}{h} \beta^{(r)} \left(\frac{x^{(s)}}{h} - k_{(m)}^{(s)} \right), \quad s = 1, \dots, d,$$

т. е. в данном случае $H_6^{(m)} = h^d$. Это означает, что при моделировании вектора $\xi_i = (\xi_i^{(1)}, \dots, \xi_i^{(d)})$ следует учесть, что компоненты $\xi_i^{(s)}$ независимы и моделируются по формуле

$$\xi_i^{(s)} = h\psi_0^{(r)} + k_{(m)}^{(s)} h, \quad (2.9)$$

где $\psi_0^{(r)}$ — выборочное значение случайной величины $\psi^{(r)}$, имеющей плотность распределения $\beta^{(r)}(z)$.

Учитывая рекуррентное определение B -сплайна r -го порядка

$$\beta^{(r)}(z) = (\beta^{(r-1)} * \beta^{(0)})(z),$$

где

$$\beta^{(0)}(z) = \begin{cases} 1 & \text{для } -1/2 \leq z \leq 1/2, \\ 0 & \text{иначе,} \end{cases} \quad (2.10)$$

а также сформулированную выше лемму и то обстоятельство, что функция $\beta^{(0)}(z)$ из (2.10) является плотностью распределения случайной величины $\alpha^{(0)} - 1/2$ ($\alpha^{(0)}$ — стандартная случайная величина), несложно обосновать моделирующую формулу

$$\psi_0^{(r)} = \left(\alpha_0 - \frac{1}{2}\right) + \left(\alpha_1 - \frac{1}{2}\right) + \dots + \left(\alpha_r - \frac{1}{2}\right) = \alpha_0 + \alpha_1 + \dots + \alpha_r - \frac{r+1}{2}, \quad (2.11)$$

где $\alpha_0, \alpha_1, \dots, \alpha_r \in U(0, 1)$ — стандартные случайные числа [3].

Что касается практического применения алгоритма 1 с формулами (2.9), (2.11), то случай $r = 0$ (кусочно-постоянное приближение (2.6)) подробно рассмотрен в работе [9].

Трудности использования “гладких” приближений (2.6), (2.8) (для достаточно больших r) описаны в диссертации [10], где подробно рассмотрен случай кубических сплайнов для $r = 3$. Они связаны, в частности, с вычислением коэффициентов $\{w_j(\mathbf{g})\}$, дающих высокий порядок аппроксимации $L_M |g(\mathbf{x})|$ (здесь может нарушаться требование положительности этих коэффициентов).

В работе [8] показано, что наиболее эффективно применение так называемого мультилинейного приближения (2.6) с базисными функциями (2.8) и сплайнами первого порядка (1.10). Особо отметим, что в практических вычислениях вместо формулы (2.11), которая для случая $r = 1$ имеет вид

$$\psi_0 = \alpha_1 + \alpha_2 - 1,$$

целесообразно использовать обоснованную в разд. 1 более экономичную формулу

$$\psi_0 = \alpha_1 - \alpha_2 \quad (2.12)$$

(см. соотношение (1.10)).

3. Стохастический алгоритм самообучения (схема Т. Кохонена) для построения адаптивных сеток в задаче приближения функций

Рассмотрим задачу численной аппроксимации функции $\varphi(\mathbf{x})$ на некоторой компактной области X из \mathbf{R}^d . Соответствующее численное приближение описывается формулой, аналогичной (2.6):

$$\varphi(\mathbf{x}) \approx L_M \varphi(\mathbf{x}) = \sum_{j=1}^M w_j(\varphi) \chi_j(\mathbf{x}). \quad (3.1)$$

Здесь, как и ранее, $\{\chi_j(\mathbf{x})\}$ — заданные базисные функции, согласованные с узлами сетки

$$X^{(M)} = \{\mathbf{x}_1, \dots, \mathbf{x}_M\}, \quad (3.2)$$

а коэффициенты $\{w_j(\varphi)\}$ зависят от значений $\varphi = (\varphi(\mathbf{x}_1), \dots, \varphi(\mathbf{x}_M))$.

Рассмотрим вопросы построения адаптивной сетки (3.2) на основе принципа эквираспределения (1.1) для задачи (3.1).

Для простоты рассмотрим сначала одномерный случай ($d = 1$), когда $Q = X = [0, 1]$. Здесь для построения приближений точек адаптивной сетки (3.2) (или (1.4)) можно применить следующую численную схему [1].

Дифференцируя соотношение (1.6) по y , получаем уравнение

$$\begin{cases} (f(x(y))x'_y(y))'_y = 0, \\ x(0) = 0, \quad x(1) = 1. \end{cases} \quad (3.3)$$

Аппроксимация этой задачи на равномерной сетке

$$\{y_j = j/M; j = 0, 1, \dots, M\} \quad (3.4)$$

в вычислительной области $Q = [0, 1]$ (см. соотношение (1.5)) приводит к системе уравнений для $x_j = x(y_j)$:

$$f_{j+1/2}x_{j+1} - (f_{j+1/2} + f_{j-1/2})x_j + f_{j-1/2}x_{j-1} = 0, \quad j = 1, \dots, M-1, \quad (3.5)$$

$$f_{j+1/2} = f(x_{j+1/2}) = f\left(\frac{x(y_j) + x(y_{j+1})}{2}\right), \quad f_{j-1/2} = f(x_{j-1/2}) = f\left(\frac{x(y_{j-1}) + x(y_j)}{2}\right),$$

которую с учетом граничных условий и того факта, что значения функции $f(x)$ заданы во всех точках $X = [0, 1]$, можно решить методом прогонки [1].

Несложно показать, что для решений $\{x_j\}$ системы (3.5) выполнен принцип эквираспределения в разностной форме (это аналог соотношения (1.8)):

$$f_{j+1/2}J_{j+1/2} = f\left(\frac{x_j + x_{j+1}}{2}\right) \frac{x_{j+1} - x_j}{h} \equiv H_7 = \text{const.} \quad (3.6)$$

Описанная методика построения приближений точек (1.4), (3.2) весьма чувствительна к размерности задачи. Когда эта размерность равна двум и более, построение соотношений типа (3.3)–(3.6) сопряжено с выполнением специальных критериев качества сетки, связанных с требованиями выпуклости и вытянутости ячеек, ортогональности линий сетки и т. д. Это обстоятельство вместе с упомянутой выше проблемой выбора управляющей функции $f(\mathbf{x})$ представляет собой трудно разрешимую проблему даже для приближения решений $\varphi(\mathbf{x})$ простейших классических задач математической физики [1].

Многих указанных недостатков лишен стохастический алгоритм построения адаптивных сеток — схема Т. Кохонена [5, 6]. Его главным преимуществом является свойство самообучения, обеспечивающее автоматизированное построение адаптивных сеток с нужными свойствами.

Кратко опишем стохастический алгоритм в общем (многомерном) случае [5, 6]. Пусть в “физической” области X (или на ее поверхности G_X) требуется построить сетку (3.2), распределение узлов которой соответствует заданной плотности $f(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}^d$. Структура этой сетки (порядок и структура расположения узлов) задается “картой” $Q^{(M)} = \{\mathbf{q}_1, \dots, \mathbf{q}_M\}$ и системой так называемых нейронов $E^{(M)} = \{\mathbf{e}_1, \dots, \mathbf{e}_M\}$ (где $\mathbf{e}_i = (\mathbf{q}_i, \mathbf{x}_i)$), определяющих соответствие между сетками $X^{(M)}$ и $Q^{(M)}$.

Приближение нейронов осуществляется с помощью процедуры самообучения, представляющей собой итерационный процесс, основанный на последовательном формировании обучающего множества $\Xi^{(T)} = \{\xi_0(1), \dots, \xi_0(T)\}$ в виде выборки из вероятностного распределения случайного вектора ξ , имеющего плотность $f(\mathbf{x})$. Здесь T — количество итераций, а $\xi_0(t) \in X$. Кроме того, с помощью значений специального коэффициента обучения $\theta_{\mathbf{q}_j}(t, \mathbf{q}_i) \in [0, 1]$ на каждом шаге итерации устанавливаются латеральные связи между нейронами \mathbf{e}_i и \mathbf{e}_j . В результате получается последовательность сеток $X^{(M)}(t) = \{\mathbf{x}_1(t), \dots, \mathbf{x}_M(t)\}$, при этом требуется выполнение соотношения

$$X^{(M)} \approx \tilde{X}^{(M)} = \lim_{t \rightarrow \infty} X^{(M)}(t) \approx X^{(M)}(T). \quad (3.7)$$

Знаки приближенных равенств в соотношении (3.7) означают не только малость расстояний $\rho(\mathbf{x}_i(\infty), \mathbf{x}_i)$, $\rho(\mathbf{x}_i(\infty), \mathbf{x}_i(T))$ и $\rho(\mathbf{x}_i(T), \mathbf{x}_i)$, где

$$\rho(\mathbf{x}, \mathbf{y}) = \sqrt{(x^{(1)} - y^{(1)})^2 + \dots + (x^{(d)} - y^{(d)})^2}, \quad \mathbf{x} = (x^{(1)}, \dots, x^{(d)}), \quad \mathbf{y} = (y^{(1)}, \dots, y^{(d)}),$$

но и воспроизведение (с хорошей точностью) требуемых свойств сетки $X^{(M)}$ (например, свойств прямоугольности, отсутствия граничного эффекта и др.) при реализации следующего итерационного процесса.

Алгоритм 2 [5, 6].

1. Устанавливаются начальные положения узлов сетки $X^{(M)}(0) = \{\mathbf{x}_1(0), \dots, \mathbf{x}_M(0)\}$.

2. На каждой итерации с номером $t = 1, \dots, T$ выполняются следующие действия:

а) выбирается очередной элемент $\xi_0(t)$ выборки $\Xi^{(T)}$;

б) вычисляются расстояния $\rho(\xi_0(t), \mathbf{x}_i(t-1))$ от точки $\xi_0(t)$ до всех узлов $\mathbf{x}_i(t-1)$ и выбирается ближайший к $\xi_0(t)$ узел $\mathbf{x}_m(t-1)$ в соответствии с условием

$$m = \arg \min_{i=1, \dots, M} \rho(\xi_0(t), \mathbf{x}_i(t-1));$$

такой узел $\mathbf{x}_m(t-1)$ называют **победителем**;

в) проводится корректировка положений всех узлов в соответствии с формулой

$$\mathbf{x}_i(t) = \mathbf{x}_i(t-1) + \theta_{\mathbf{q}_m}(t, \mathbf{q}_i) (\xi_0(t) - \mathbf{x}_i(t-1)) \quad (3.8)$$

для всех $i = 1, \dots, M$.

На каждой итерации (3.8) основного алгоритма узлы сетки сдвигаются к случайной точке $\xi_0(t)$. Поэтому в местах высокой концентрации элементов выборки постепенно скапливается все больше узлов, за счет чего достигаются сгущения сетки. В работе [6] показано, что при $T \rightarrow \infty$ и специальном выборе коэффициента обучения $\theta_{\mathbf{q}_j}(t, \mathbf{q}_i)$ основной алгоритм ведет к выполнению принципа эквираспределения, т.е. к получению требуемой плотности сетки, определяемой функцией $f(\mathbf{x})$ (см. соотношения (1.1), (1.7)). Отметим также, что в работах [11, 12] предложены конструктивные подходы к аналитическому исследованию алгоритма 2.

Интересен и принципиален вопрос о выборе плотности $f(\mathbf{x})$. Здесь определенные трудности возникают даже в одномерном случае. Например, в пособии [1] без особой аргументации для достаточно ограниченного класса задач предлагается выбирать управляющую функцию в виде

$$f(x) = 1 + C_0 |\varphi(x)|^{b_0} + C_1 |\varphi'(x)|^{b_1} + C_2 |\varphi''(x)|^{b_2},$$

где неотрицательные константы C_k, b_k , $k = 0, 1, 2$, подбираются экспериментально, а вместо производных берутся их сеточные приближения.

Авторами статьи предлагается следующий конструктивный подход к выбору плотности $f(x)$. Рассмотрим кусочно-линейное приближение на равномерной сетке $\{y_i\}$ из соотношения (3.4):

$$L_M \varphi(x) = \varphi(y_j) + (x - y_j) \frac{\varphi(y_{j+1}) - \varphi(y_j)}{y_{j+1} - y_j}, \quad j = 0, 1, \dots, M-1. \quad (3.9)$$

Используя рассуждения из параграфа 3 гл. 2 монографии [4], можно показать, что для этого приближения выполнено соотношение

$$\rho_{C[0,1]}(\varphi, L_M \varphi) = \sup_{x \in [0,1]} |\varphi(x) - L_M \varphi(x)| \leq \frac{1}{8} \sup_{x \in [0,1]} |\varphi''(x)| h^2, \quad (3.10)$$

$$h = y_{j+1} - y_j = 1/M, \quad j = 0, 1, \dots, M-1.$$

Введем допустимый уровень погрешности δ и приравняем его правой части неравенства (3.10):

$$\frac{1}{8} \sup_{x \in [0,1]} |\varphi''(x)| h^2 = \delta \quad \text{или} \quad \sqrt{\sup_{x \in [0,1]} |\varphi''(x)|} \cdot h = H_8, \quad (3.11)$$

где $H_8 = \sqrt{8\delta}$. Сравнительный анализ соотношений (1.8) и (3.11) приводит к рекомендации использовать в качестве плотности $f(x)$ сеточное приближение функции

$$\tilde{f}(x) = H_9 \sqrt{|\varphi''(x)|}, \quad (3.12)$$

где H_9 — соответствующая нормировочная константа.

Приведем пример, показывающий целесообразность выбора плотности в соответствии с рекомендацией (3.12). Рассмотрим функцию

$$\varphi(x) = e^{-2x}, \quad x \in [0, 2].$$

Введем равномерную сетку $x_j^{(uniform)} = (j-1)\Delta$, $j = 1, \dots, 9$, с шагом $\Delta = 0.25$. Затем, следуя рекомендации (3.12), определим плотность $f(x)$ в виде

$$f(x) = \frac{e^2}{e^2 - 1} e^{-x}, \quad x \in [0, 2]. \quad (3.13)$$

Здесь использовано обстоятельство, что $\varphi''(x) = 4e^{-2x}$.

Заметим, что для плотности (3.13) можно построить моделирующую формулу метода обратной функции распределения [3] для соответствующего выборочного значения

$$\xi_0 = -\ln \left(1 - \frac{e^2 - 1}{e^2} \alpha_0 \right),$$

где $\alpha_0 \in U(0, 1)$ — стандартное случайное число. Эта формула может быть применена при реализации соответствующей одномерной версии алгоритма 2.

Применим принцип эквираспределения и рассмотрим соответствующее соотношение (1.6) с граничными условиями

$$f(x(y))x'(y) = H_{10}, \quad H_{10} = \text{const},$$

$$x(0) = 0, \quad x(1) = 2.$$

Решаем получившееся дифференциальное уравнение и находим $x(y)$. Общее решение имеет вид

$$\left(\frac{e^{2-x}}{e^2 - 1} \right) dx = H_{10} dy$$

или

$$-e^{2-x} = H_{10}(e^2 - 1)y + H_{11}, \quad H_{11} = \text{const}.$$

Определяем константы из граничных условий

$$H_{10} = 1, \quad H_{11} = -e^2.$$

Окончательно получаем

$$x(y) = 2 - \ln(e^2 - (e^2 - 1)y), \quad y \in [0, 1].$$

Результаты пересчета узлов адаптивной сетки

$$x_j^{(adaptive)} = x((j - 1)h), \quad j = 1, \dots, 9, \quad h = 0.125,$$

запишем в табличной форме. Здесь $x_j^{(uniform)}$ — узлы равномерной сетки, $x_j^{(adaptive)}$ — узлы адаптивной сетки.

Далее построим кусочно-линейное приближение функции $\varphi(x) = e^{-2x}$ на равномерной и адаптивной сетках. Для этого вычислим значения этой функции в узлах равномерной и адаптивной сеток.

$x_j^{(uniform)}$	0	0.25	0.5	0.75	1	1.25	1.5	1.75	2
$x_j^{(adaptive)}$	0	0.11	0.24	0.39	0.57	0.78	1.05	1.41	2

$e^{-2x_j^{(uniform)}}$	1	0.6	0.368	0.223	0.135	0.082	0.05	0.03	0.018
$e^{-2x_j^{(adaptive)}}$	1	0.8	0.618	0.458	0.320	0.210	0.122	0.06	0.018

Найдем погрешности

$$\delta^{(adaptive)} = \rho_{C[0,2]}(\varphi, L_M^{(adaptive)}\varphi) = \max_{x \in [0,2]} |\varphi(x) - L_M^{(adaptive)}\varphi(x)|,$$

$$\delta^{(uniform)} = \rho_{C[0,2]}(\varphi, L_M^{(uniform)}\varphi) = \max_{x \in [0,2]} |\varphi(x) - L_M^{(uniform)}\varphi(x)|,$$

где $L_M^{(adaptive)}\varphi(x)$ — кусочно-линейная аппроксимация (3.9) с узлами $\{x_j^{(adaptive)}\}$, а $L_M^{(uniform)}\varphi(x)$ — кусочно-линейная аппроксимация (3.9) с узлами $\{x_j^{(uniform)}\}$.

На равномерной сетке максимальное отклонение $\delta^{(adaptive)}$ достигается на интервале $[0, 0.25]$ и равно 0.024. Соответствующие вычисления на адаптивной сетке дают максимальное отклонение $\delta^{(uniform)}$ на интервале $[0.57, 0.78]$, равное 0.01. Получаем, что погрешность аппроксимации на адаптивной сетке заметно меньше.

Отметим, что если функция $\varphi(x)$ не задана явно, а лишь известны ее значения в узлах сетки, то для вычисления второй производной в узлах сетки можно использовать ее сеточные приближения (по аналогии с работой [1]).

Заключение

Перечислим основные результаты данной работы.

1. Получен вероятностный аналог принципа эквираспределения (в интегральной форме, одномерный случай) — см. соотношения (1.9), (1.10).

2. Показано, что метод выборки по важности, широко используемый в задачах многомерного численного стохастического интегрирования, связан с формированием сетки (2.5), адаптивной в том смысле, что плотность распределения узлов пропорциональна модулю подынтегральной функции. Приведен также конструктивный алгоритм метода

суперпозиции (алгоритм 1) с конкретным аппроксимационным базисом (2.8) и эффективными (экономичными) моделирующими формулами (2.9), (2.12).

3. Для задачи приближения функций многих переменных рассмотрен стохастический итерационный алгоритм построения адаптивных сеток (схема Т. Кохонена — алгоритм 2). Предложен подход к выбору плотности распределения адаптивных узлов (одномерный случай — соотношение (3.12)) и на простом примере показана эффективность такого подхода.

Отметим следующие возможности развития и применения перечисленных результатов.

1. Относительно несложно получить многомерные аналоги полученных здесь результатов (например, соотношений (1.9), (1.10) и (3.12)). В частности, индукцией по размерности d , используя разложение совместной плотности распределения случайного вектора в произведение условных плотностей [2, 3], можно получить следующий аналог формулы (1.9) для одинаково распределенных (согласно плотности $f(\mathbf{z}) = f(z^{(1)}, \dots, z^{(d)})$) d -мерных случайных векторов $\xi_1 = (\xi_1^{(1)}, \dots, \xi_1^{(d)})$ и $\xi_2 = (\xi_2^{(1)}, \dots, \xi_2^{(d)})$:

$$\int_{\xi_1^{(1)}}^{\xi_2^{(1)}} \dots \int_{\xi_1^{(d)}}^{\xi_2^{(d)}} f(z^{(1)}, \dots, z^{(d)}) dz^{(1)} \dots dz^{(d)} = \alpha^{(1)} \cdot \dots \cdot \alpha^{(d)} - \alpha^{(d+1)} \cdot \dots \cdot \alpha^{(2d)}, \quad (4.1)$$

где $\alpha^{(s)}$, $s = 1, \dots, 2d$, — независимые стандартные (равномерно распределенные в интервале $(0, 1)$) случайные величины.

Заметим, что формула (4.1) и ее обоснование подтверждают высказанный во введении статьи тезис о том, что многомерные аналоги полученных в данной работе результатов и их обоснований имеют относительно простой, но достаточно громоздкий вид. Такой вид будет иметь и многомерный аналог формулы (1.10). Здесь для получения плотности распределения случайной величины из правой части соотношения (4.1) следует использовать известные формулы для произведения и суммы независимых случайных величин с известными плотностями распределения [2].

2. Для развития теории и приложений конструктивных алгоритмов выборки по важности целесообразны дополнительные исследования моделируемых базисов, отличных от конечных элементов (2.8). Особую роль здесь играет построение экономичных реализаций соответствующих алгоритмов с целью их включения в геометрические вычислительные модули (подобно тому, как это сделано в работе [9]).

3. Возможно обобщение подхода к выбору плотности узлов в задаче аппроксимации функций из разд. 3 на многомерный случай и (или) на случай использования более гладких аппроксимаций $L_M \varphi(\mathbf{x})$ (см. соотношение (3.1)). Здесь можно предложить (по аналогии с формулой (3.12)) конструирование управляющей функции (плотности) $f(\mathbf{x})$ в виде комбинации модулей производных приближаемой функции, входящих в выражения для оценок погрешностей $\rho_{\mathbf{B}(X)}(\varphi, L_M \varphi) = \|\varphi(\mathbf{x}) - L_M \varphi(\mathbf{x})\|_{\mathbf{B}(X)}$ в норме соответствующего функционального пространства $\mathbf{B}(X)$ [3, 4, 7, 10]. Важно также умножать примеры эффективного использования многомерных адаптивных сеток при решении прикладных задач (подобно тому, как это сделано в пособии [1]).

Список литературы / References

- [1] **Хакимзянов Г.С., Шокин Ю.И.** Разностные схемы на адаптивных сетках. Ч. 1, 2. Новосибирск: НГУ, 2009.
Khakimzyanov, G.S., Shokin, Yu.I. Difference schemes on adaptive meshes. Vol. 1, 2. Novosibirsk: NGU, 2009 (In Russ.)
- [2] **Боровков А.А.** Теория вероятностей. М.: Наука, 1986. 432 с.
Borovkov, A.A. Probability theory. Moscow: Nauka, 1986. 432 p. (In Russ.)
- [3] **Михайлов Г.А., Войтишек А.В.** Численное статистическое моделирование. Методы Монте-Карло. М.: Изд. центр “Академия”, 2006. 368 с.
Mikhaylov, G.A., Voytishkek, A.V. Numerical statistical modelling. Monte Carlo methods. Moscow: Izd. Tsentr “Akademiya”, 2006. 368 p. (In Russ.)
- [4] **Березин И.С., Жидков Н.П.** Методы вычислений. Т. 1. М.: Гос. изд-во физ.-мат. лит. 1959. 464 с.
Berezin, I.S., Zhidkov, N.P. Numerical methods. Vol. 1. Moscow: Gos. izd-vo fiz.-mat. lit., 1959. 464 p. (In Russ.)
- [5] **Kohonen, T.** Self-organizing maps. Springer-Verlag, 2001. 502 p.
- [6] **Нечаева О.И.** Нейросетевые модели, алгоритмы и комплекс программ для построения адаптивных сеток: Дис. ... канд. физ.-мат. наук. Новосибирск: НГУ, 2007. 120 с.
Nechaeva, O.I. Neural network models, algorithms and program complex for constructing adaptive meshes: Dis. ... kand. fiz.-mat. nauk. Novosibirsk: NGU, 2007. 120 p. (In Russ.)
- [7] **Марчук Г.И., Агошков В.И.** Введение в проекционно-сеточные методы. М.: Наука, 1981. 416 с.
Marchuk, G.I., Agoshkov, V.I. Introduction in projective-mesh methods. Moscow: Nauka, 1981. 416 p. (In Russ.)
- [8] **Voytishkek, A.V., Kablukova, E.G.** Using the approximation functional bases in Monte Carlo methods // Russ. J. of Numer. Analysis and Mathem. Modelling. 2003. Vol. 18, No. 6. P. 521–542.
- [9] **Бессмельцев М.В., Войтишек А.В.** Модификации геометрических программных модулей, связанные с построением моделируемых вероятностных плотностей // Вычисл. технологии. 2011. Т. 16, № 4. С. 19–36.
Bessmeltsev, M.V., Voytishkek, A.V. Modification of geometrical program modules, tied with construction of modelled probabilistic densities // Comput. Technologies. 2011. Vol. 16, No. 4. P. 19–36. (In Russ.)
- [10] **Милосердов В.В.** Дискретно-стохастические численные алгоритмы со сплайн-восполнениями: Дис. ... канд. физ.-матем. наук. Новосибирск: НГУ, 2006. 83 с.
Miloserdov, V.V. Discrete-stochastic numerical algorithms with spline approximations: Dis. ... kand. fiz.-mat. nauk. Novosibirsk: NGU, 2007. 83 p. (In Russ.)
- [11] **Войтишек А.В., Хмель Д.С.** Аналитическое описание применения одномерной схемы Т. Кохонена для построения адаптивных сеток // Сиб. журн. вычисл. математики. 2011. Т. 14, № 2. С. 131–140.
Voytishkek, A.V., Khmel, D.S. Analytical description for application of the 1D Kohonen scheme for constructing adaptive grids // Numerical Analysis and Applications. 2011. Vol. 4, No. 2. P. 105–113.

- [12] **Войтишек А.В., Хмель Д.С.** Аналитический подход к изучению граничного эффекта в одномерном рандомизированном численном алгоритме построения адаптивных сеток // Журн. вычисл. математики и матем. физики. 2013. Т. 53, № 2. С. 195–208.
Voytishek, A.V., Khmel, D.S. Analytical approach to study the boundary effect in the 1D randomized numerical algorithm for constructing adaptive grids // Zurn. Vychisl. Matematiki i Matem. Fiziki. 2013. Vol. 53, No. 2. P. 195–208. (In Russ.)

*Поступила в редакцию 24 августа 2016 г.,
с доработки – 13 января 2017 г.*

On the choice of distribution densities in adaptive mesh nodes for stochastic algorithms of numerical integration and function approximation

VOYTISHEK, ANTON V.*, PRASOL, DENIS A.

Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, 630090, Russia

*Corresponding author: Voytishek, Anton V., e-mail: vav@osmf.ssc.ru

In this paper we formulate the analog of the equidistributional principle for stochastic adaptive meshes which are used for solving the multi-dimensional problems of numerical integration. It means that the integral for density of continuous 1D random variable ξ over the interval (ξ_1, ξ_2) between two sample values of the variable ξ has the “universal” distribution with the density equal to the so called “hat-function”.

The importance sampling technique is treated as a method for constructing an adaptive mesh for the problem of numerical integration. The discrete-stochastic approach with the “modelled” basis of the Strang – Fix approximation for numerical simulation of the corresponding adaptive mesh is proposed.

We also consider the problem on the choice of the distribution density of nodes in implementation of self-organization in the T. Kohonen’s algorithm. The research is aimed at construction of adaptive meshes for solving the problem of numerical function approximation. For the smooth approximated functions we propose to use the square root of the second derivative module of the approximated function as density of adaptive nodes. The example of function for which the formulated recommendation leads to improvement of approximating properties of mesh approach is given.

The most part of results is formulated in the simplest 1D-version. The survey of possible multi-dimensional versions of these results and directions of further investigations is presented in the conclusion.

Keywords: adaptive mesh, equidistributional principle, probabilistic distribution density, Monte Carlo method, importance sampling method, T. Kohonen’s scheme, choice of nodes density for function approximation.

Received 24 August 2016

Received in revised form 13 January 2017