

УДК 539.122:518.5

## ФУНКЦИОНАЛЬНЫЕ ОЦЕНКИ В МЕТОДАХ СТАТИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ

Д. Г. Модестов, К. Е. Хатунцев  
(РФЯЦ-ВНИИТФ)

Предлагается способ оценки распределений, которые индуцируются вероятностной мерой, используемой при решении задач методами статистического моделирования. Описывается критерий приближения, приводятся явные выражения для оценки его погрешности.

*Ключевые слова:* методы статистического моделирования, математическое ожидание, ковариационная матрица, дисперсия, производная Радона—Никодима, плотность заряда, приближение функций.

### Введение

Как известно (см., например, [1, 2]), методы статистического моделирования основаны на получении оценки математического ожидания случайных величин. Сами эти величины являются измеримыми функциями на вероятностном пространстве, которое используется для построения процедуры выборки, а математическое ожидание — интегралом по этому пространству [2]. Само вероятностное пространство может быть достаточно велико, например, бесконечномерное пространство траекторий марковского процесса. Однако на практике часто интересны случайные величины, принадлежащие некоторому ограниченному классу, математические ожидания которых могут быть представлены интегралом соответствующих функций по гораздо меньшему пространству с некоторой, не обязательно вероятностной, мерой или зарядом (неположительной мерой — см., например, [3]). Например, при решении задачи переноса путем имитации [1, 2] математическими ожиданиями, в частности, являются число частиц в некоторой области в заданный момент времени, энергия этих частиц, число реакций определенного типа в единицу времени в этой же области. Для представления этих значений в интегральном виде обычно на множестве значений энергии определяют функцию, называемую спектром или дифференциальной по энергии плотностью числа частиц, которая является некоторым решением уравне-

ния переноса. Следует отметить, что данная функция в общем случае является обобщенной и может иметь сингулярную часть, обычно связанную с процессами рождения частиц при релаксации возбужденных состояний либо с наличием непроизводимых частиц. Указанные выше математические ожидания вычисляются интегрированием по спектру единичной функции (т. е. функции, всюду принимающей значение единица), энергии и произведения скорости на сечение реакции соответственно.

Таким образом, спектр определяет некоторую меру на множестве значений энергии. Он сам может также представлять интерес — для вычисления математического ожидания функций, которые по каким-либо причинам не могли быть использованы во время решения задачи, анализа состояния системы, для представления и сравнения результатов расчета, для передачи решения на следующий этап, если производится многоэтапный расчет, и т. д. Но так как в общем случае спектр может быть локализован только на пространстве интегрируемых функций, являющимся бесконечномерным, его представление может быть только приближенным. Соответственно возникает задача выбора приемлемого приближения.

Таким образом, возвращаясь к общему случаю, можно сказать, что при решении прикладных задач методами статистического моделирования возникает подзадача представления на

некотором множестве параметров заряда, индуцированного вероятностной мерой самой задачи.<sup>1</sup>

Наиболее часто для этого используется представление плотности этого заряда (спектра в приведенном примере) в виде кусочно-постоянной функции с конечным числом значений, которую принято называть гистограммой. Можно показать (см., например, [3]), что если мера абсолютно непрерывна относительно меры Лебега, то ее можно приблизить гистограммой с любой точностью. То же самое утверждение легко экстраполировать и на заряд. Кроме того, для вычисления интегралов, необходимых для построения гистограммы с произвольным разбиением, несложно построить эффективный алгоритм суммирования выборочных значений. Поэтому такое представление является в каком-то смысле универсальным и соответственно наиболее часто используемым. Кроме того, оно является вполне удобным для анализа, так как в достаточном числе случаев по результатам расчета, без предварительной обработки, можно выявить некоторые особенности решения. Скажем, в приведенном выше примере, в случае, если для гистограммы задано равномерное разбиение, приближенно определить положение максимума спектра можно, просматривая оценки интегралов.

Однако во многих случаях плотность заряда является функцией высокой степени гладкости. Следовательно, можно ожидать, что ее представление с помощью гладких функций может быть предпочтительнее. Некоторые аспекты использования для этой цели системы ортогональных функций обсуждаются в работах [2, 4]. Пример использования ортогональной системы многочленов приводится в [5]. Следует заметить, что и для построения гистограммы фактически используется система ортогональных функций. Для использования таких систем необходимо предварительное, т. е. до начала расчета задачи, знание особенностей поведения плотности меры. Например, для гистограммы необходимо знать области, в которых происходит быстрое изменение плотности, так как в них требуется более подробное разбиение этой гистограммы. Для использования же ортогональных многочленов необходимо знать, что оцениваемая плотность хорошо приближается данной систе-

мой. Обычно, когда идет расчет больших серий однотипных задач, такое знание приобретается проведением нескольких настроечных расчетов. Но если для использования гистограмм такой метод достаточно эффективен, то для ортогональных многочленов его использование несколько затруднено. Помимо того, что далеко не все программы, использующие метод статистического моделирования, имеют достаточно гибкую систему ввода данных, позволяющую задавать произвольные оценки, это затруднение вызвано сложностью подбора подходящей ортогональной системы во время проведения настроечных расчетов, особенно если эти расчеты трудоемки.

Однако если отказаться от требования обязательной ортогональности, то методика становится более гибкой. При этом заметно большее значение приобретает пострасчетная обработка. Данное обстоятельство, впрочем, на взгляд авторов, при наличии персональной ЭВМ может привести, скорее, к психологическим затруднениям в использовании методики. Следует также заметить, что одной из сильных сторон методов статистического моделирования является оценка погрешности рассчитываемых величин, получаемая одновременно с оценкой значений этих величин. Поэтому кажется вполне естественным, что методика оценки плотности заряда должна содержать и оценку статистической погрешности представления, причем эта оценка должна быть как локальной (например, погрешность в точке непрерывности), так и интегральной.

Учитывая эти обстоятельства, в настоящей работе делается попытка описать непротиворечивую методику оценки распределений.

### Критерий приближения

Таким образом, постановка подзадачи представления заряда является следующей. Пусть задано множество значений параметров  $X$  с  $\sigma$ -алгеброй множеств на нем (определение алгебры см. в [3]), которая однозначно определяется постановкой вероятностной задачи. В абсолютном большинстве практических случаев  $X$  является подмножеством конечномерного евклидова пространства, а  $\sigma$ -алгеброй является система борелевских множеств, и потому всюду в дальнейшем именно она и будет подразумеваться по умолчанию (хотя такое ограничение не принципиально). Требуется найти приближенное представление заряда, индуцированного вероятностной мерой решаемой задачи.

<sup>1</sup>Чаще всего интересны меры, но для общности будет использоваться понятие заряда.

Строгое определение индуцированного заряда, который в настоящей работе будет обозначаться символом  $\mu$ , дается значениями интеграла для любой измеримой функции. Для этого удобно обозначить через  $\eta(f)$  случайную величину, используемую при оценке вклада функции  $f$ . Для пояснения можно сказать, что в примере, приведенном во Введении, при решении задачи аналоговым методом эта величина определяется как

$$\eta(f) = \sum_k f(E_k),$$

где  $E_k$  — энергии всех частиц, которые на одном статистическом испытании при пересечении заданной временной точки находились в заданной пространственной области. При использовании весовых методов значение случайной величины будет следующим:

$$\eta(f) = \sum_k w_k f(E_k),$$

где дополнительно через  $w_k$  обозначены статистические веса этих частиц. С использованием данного обозначения интеграл функции по заряду определяется как математическое ожидание соответствующей случайной величины:

$$\int f(x) \mu(dx) = M\eta(f). \quad (1)$$

Следует отметить, что если для любой неотрицательной функции случайная величина  $\eta$  также неотрицательна, как это, в частности, справедливо для приведенного примера, то  $\mu$  является мерой.

Для решения подзадачи представления необходимо выделить некоторый класс мер, элементы которого будут называться базовыми мерами и обозначаться как  $\mu_0$ . Одним из условий принадлежности меры к этому классу является то, что заряд  $\mu$  абсолютно непрерывен относительно  $\mu_0$ , т. е. для всякого измеримого множества  $Y \subset X$  справедливо  $\mu_0(Y) = 0 \Rightarrow \mu(Y) = 0$ . Как следует из теоремы Радона—Никодима [3], в этом случае существует  $\mu_0$ -измеримая функция  $\rho \equiv \frac{d\mu}{d\mu_0}$  такая, что для любой  $\mu_0$ -измеримой (и интегрируемой) функции  $f$  справедливо равенство

$$\int f(x) \mu(dx) = \int \rho(x) f(x) \mu_0(dx). \quad (2)$$

Функция  $\rho$  здесь для краткости будет называться плотностью заряда.<sup>2</sup> Вторым условием принадлежности  $\mu_0$  к классу базовых мер является то, что функция  $\rho$  является квадратично-интегрируемой по этой мере:  $\rho \in L_2(X, \mu_0)$ . В практически важных случаях обычно несложно выделить непустое подмножество мер, удовлетворяющих данному условию.

Таким образом, представление меры эквивалентно приближенному представлению плотности при подходящем выборе базовой меры. Для этого будет использоваться линейная комбинация конечного набора независимых функций  $\varphi_k$

$$\rho(x) \approx \rho_n(x) = \sum_{k=1}^n c_k \varphi_k(x), \quad (3)$$

линейная оболочка которых будет называться пространством представления ( $n$  — его размерность). Естественным ограничением на сами функции является то, что они должны быть интегрируемыми по заряду. Для того чтобы задача была однозначно разрешимой, эти функции должны быть линейно независимы на носителе  $\mu_0$  (т. е. минимальном подмножестве, дополнение к которому имеет нулевую меру [3]). Кроме того, так как для практики наиболее интересной задачей является приближение действительной функции, то  $\varphi_k$  и соответственно  $c_k$  естественно выбирать действительными, что и будет предполагаться в дальнейшем. Впрочем, обобщение на комплексно-значимые функции очевидно. Также, с учетом уже сделанного ограничения на плотность, эти функции должны быть квадратично-интегрируемыми по базовой мере. Предполагается, что сама система функций является частью полной системы, т. е. при увеличении  $n$  плотность может быть представлена сколь угодно точно:

$$\lim_{n \rightarrow \infty} \rho_n(x) = \rho(x).$$

Удобно рассмотренные ограничения представить совместно:

- 1)  $\varphi_k|_{k=1}^n$  такие, что  $|\int \varphi_k(x) \mu(dx)| < \infty$ ;
- 2)  $\mu_0$  такая, что  $\frac{d\mu}{d\mu_0}$ ,  $\varphi_k \in L_2(X, \mu_0)$  и  $\varphi_k$  линейно независимы на носителе  $\mu_0$ .

Следует заметить, что ограничение 2 неявно содержит существование производной Радона—Никодима

<sup>2</sup>Наиболее распространенное название — производная Радона—Никодима  $\mu$  по  $\mu_0$ , но для описания оно представляется слишком громоздким.

кодима и, следовательно, требование абсолютной непрерывности.

Для большинства практических задач подобные ограничения не являются слишком обременительными. Скажем, для приведенного во Введении примера при условии отсутствия у спектра сингулярной части им удовлетворяют степенные функции и  $\mu_0(dE) = \rho_0(E)dE$ , где  $\rho_0$  — произвольная строго положительная, быстро убывающая функция.

Ограничение 2 было сделано для того, чтобы можно было удобным образом определить понятие близости. Соответственно в качестве условия приближения будет использоваться минимизация расстояния между функциями в пространстве  $L_2(X, \mu_0)$  с учетом того, что

$$\begin{aligned} \Delta_n &= \min_{c_1, \dots, c_k} \|\rho - \rho_n\|_{\mu_0} = \\ &= \min_{c_1, \dots, c_k} \left\| \rho - \sum_{k=1}^n c_k \varphi_k \right\|_{\mu_0}. \end{aligned} \quad (4)$$

Здесь символом  $\|f\|_{\mu_0}$  обозначена норма в пространстве  $L_2(X, \mu_0)$  (см., например, [3]):

$$\|f\|_{\mu_0}^2 = \int |f(x)|^2 \mu_0(dx).$$

Несложно видеть, что это условие эквивалентно следующей системе линейных уравнений:

$$\sum_{k=1}^n c_k \int \varphi_m(x) \varphi_k(x) \mu_0(dx) = \int \varphi_m(x) \mu(dx),$$

$$m = 1, \dots, n.$$

Данную систему удобно записать в матричном виде. Для упрощения записи при этом большими буквами будут обозначаться квадратные матрицы, малыми — вектор-столбцы, значком "+" в верхнем индексе — операция транспонирования. Например,  $\mathbf{c}$  — вектор-столбец с компонентами  $c_k$ ;  $\mathbf{c}^+$  — вектор-строка с этими же компонентами. Таким образом, минимизирующая система уравнений имеет вид

$$A\mathbf{c} = \mathbf{b}, \quad (5)$$

где  $A$  — симметричная матрица с компонентами  $A_{m,k} = \int \varphi_m(x) \varphi_k(x) \mu_0(dx)$ ;  $\mathbf{b}$  — вектор-столбец с компонентами  $b_m = \int \varphi_m(x) \mu(dx)$ . Можно показать, что матрица  $A$  является положительно определенной и, следовательно, обратная матрица существует, а система (5) всегда имеет единственное решение.

Несложно показать, что задача приближения зависит только от выбора линейного пространства, в котором должно лежать  $\rho_n$ , и не зависит от выбора конкретного базиса. Это следует уже из выражения (3), которое инвариантно по отношению к замене базиса. Однако для получения явных выражений удобно рассмотреть это свойство на решениях уравнения (5).

Если на линейной оболочке набора  $\varphi_k$  выбрать другую линейно независимую систему функций  $\tilde{\varphi}_m$ , связанную с исходной системой матрицей  $V$ ,

$$\tilde{\varphi}_m(x) = \sum_{k=1}^n V_{m,k} \varphi_k(x), \quad (6)$$

то легко проверить, что связь между объектами, используемыми в уравнении (5), в разных системах будет иметь вид (волнистая черта сверху означает, что объекты вычислены в системе  $\tilde{\varphi}_m$ )

$$\tilde{\mathbf{b}} = V\mathbf{b}, \quad \tilde{A} = VAV^+, \quad A^{-1} = V^+ \tilde{A}^{-1} V. \quad (7)$$

Следует отметить, что вследствие линейной независимости функций  $\tilde{\varphi}_m$  матрица  $V$  является обратной. Решение минимизирующего уравнения  $\tilde{A}\tilde{\mathbf{c}} = \tilde{\mathbf{b}}$ , как можно видеть, будет следующим:

$$\tilde{\mathbf{c}} = \tilde{A}^{-1} \tilde{\mathbf{b}} = (V^{-1})^+ \mathbf{c}.$$

Явный вид плотности заряда получается подстановкой этого решения в выражение (3) с учетом (6):

$$\begin{aligned} \tilde{\rho}_n(x) &= \sum_{k=1}^n \tilde{c}_k \tilde{\varphi}_k(x) = \sum_{k,m,p=1}^n V_{m,k}^{-1} c_m V_{k,p} \varphi_p(x) = \\ &= \sum_{m=1}^n c_m \varphi_m(x) = \rho_n(x), \end{aligned}$$

где через  $V_{m,k}^{-1}$  обозначены элементы обратной матрицы. Из совпадения этих функций и следует инвариантность задачи приближения относительно замены базиса.

Таким образом, система (5) решает подзадачу приближения заряда. При этом компоненты вектора, стоящего в правой части этой системы, в соответствии с (1) являются математическими ожиданиями соответствующих случайных величин:

$$b_m = M\eta(\varphi_m). \quad (8)$$

Значения этих математических ожиданий оцениваются во время решения основной задачи.

Здесь следует отметить то обстоятельство, что во время основного расчета, который может быть весьма продолжительным, оцениваются только величины (8). При этом применяется только набор функций, используемых для приближения плотности в выражении (3). Базовая мера при этом никак не используется. Следовательно, возникает возможность выбрать ее таким образом, чтобы приближенное значение плотности достаточно хорошо представляло особенности индуцированного заряда. Однако сначала следует рассмотреть погрешность представления.

### Оценка погрешности

Как можно видеть, существуют две погрешности представления различной природы: обусловленные конечностью линейного пространства и статистическим разбросом.

Численным значением первой погрешности является  $\Delta_n$  — расстояние в пространстве  $L_2(X, \mu_0)$  между плотностью заряда и линейной оболочкой функций  $\varphi_k$ , явный вид которого при условии, что коэффициенты  $c_k$  удовлетворяют уравнению (5), дается выражением (4). К сожалению, не представляется возможным сделать строгую оценку этой величины в общем случае. Можно лишь указать следующие свойства:

- 1)  $n_1 < n_2 \Rightarrow \Delta_{n_1} \geq \Delta_{n_2}$ ;
- 2)  $\Delta_n \xrightarrow{n \rightarrow \infty} 0$ .

Второе свойство связано с тем, что система функций  $\varphi_k$  по предположению является частью полной системы в пространстве  $L_2(X, \mu_0)$ . Таким образом, можно утверждать, что с ростом размерности представления погрешность монотонно убывает к нулю. В некоторых частных случаях можно дать более строгие оценки, однако рассмотрение этих случаев выходит за рамки настоящей работы.

Статистическая погрешность возникает из-за того, что оценки математических ожиданий (8), полученные как результаты моделирования конечного числа испытаний, лишь приближенно совпадают с истинными величинами. Для оценки погрешности математического ожидания некоторой случайной величины  $\eta$  при расчете ее методами статистического моделирования обычно используется среднее квадратичное отклонение этой величины

$$\sigma(\eta) = \sqrt{\frac{D\eta}{N}}, \quad (9)$$

где  $D\eta$  — дисперсия случайной величины  $\eta$  [1, 2];  $N$  — число испытаний, произведенных для оценки. Здесь предполагается, что дисперсия конечна. В дальнейшем это предположение будет считаться справедливым для всех случайных величин, используемых в (8). Подробнее о связи среднего квадратичного отклонения с погрешностью метода можно узнать из книг [1, 2], здесь достаточно сказать, что с определенной вероятностью этой величине пропорциональна разность между истинным и выборочным значением математического ожидания. Таким образом, для обеспечения возможности оценки погрешности достаточно определить значение дисперсии рассчитываемой величины.

Как уже упоминалось, наиболее интересными с прикладной точки зрения функционалами являются интеграл некоторой функции по заряду и значение плотности заряда в точке. Последнее необходимо, в частности, для представления этой плотности. Что касается оценки интеграла, то его значение для некоторой функции  $f$  определяется проекцией этой функции на линейную оболочку  $\varphi_k|_{k=1}^n$ . Это следует из того, что интеграл от произведения плотности (3) и любой функции, перпендикулярной в смысле  $L_2(X, \mu_0)$  этой оболочке, тождественно равен нулю.

Для нахождения дисперсии удобно обозначить

$$f_k = \int f(x) \varphi_k(x) \mu_0(dx); \quad \eta_k = \eta(\varphi_k). \quad (10)$$

Естественным ограничением на функции, для которых возможна подобная оценка, является существование данного интеграла. В частности, это всегда справедливо для  $f \in L_2(X, \mu_0)$ .

С использованием этих обозначений случайная величина, применяемая для оценки функционала  $\int \rho_n(x) f(x) \mu_0(dx)$ , с учетом (3), (5), (8) и (10) имеет вид

$$\zeta_f = \sum_{k,m=1}^n f_k A_{k,m}^{-1} \eta_m, \quad (11)$$

где через  $A_{k,m}^{-1}$  обозначены элементы обратной матрицы. Как можно видеть, необходимое для вычисления дисперсии математическое ожидание ее квадрата может быть представлено следующим выражением:

$$M\zeta_f^2 = \sum_{k,m,q,p=1}^n f_k A_{k,m}^{-1} f_q A_{q,p}^{-1} M(\eta_m \eta_p).$$

Для представления явного выражения дисперсии удобно обозначить

$$K_{m,p} = M(\eta_m \eta_p) - M\eta_m M\eta_p. \quad (12)$$

Матрицу, элементами которой являются эти величины, принято называть ковариационной матрицей [2]. Сами значения (12) оцениваются во время расчета методом, аналогичным эмпирической оценке дисперсии [1]. Несложно показать, используя неравенство Коши—Буняковского [3], что существование дисперсий всех величин  $\eta_k$  является достаточным условием существования ковариационной матрицы и, следовательно, при введенных ограничениях на дисперсию определение этой матрицы корректно. Кроме того, можно показать, что эта матрица неотрицательна.

С использованием указанных обозначений, а также учитывая представление дисперсии как  $D\zeta = M\zeta^2 - (M\zeta)^2$  (см., например, [1, 2]), дисперсию случайной величины (11) можно представить в матричном виде:

$$D\zeta_f = f^+ A^{-1} K A^{-1} f, \quad (13)$$

где явно учтена симметричность матрицы  $A$ .

В некоторых случаях удобно использовать альтернативное представление дисперсии. Для этого можно определить симметричную функцию двух аргументов, которая будет называться ковариационной функцией:

$$K(x, y) = \sum_{k,m=1}^n K_{k,m}^A \varphi_k(x) \varphi_m(y), \quad (14)$$

где  $K^A = A^{-1} K A^{-1}$  (использование одного символа для обозначения матрицы и функции показывает, что это фактически представления одного оператора на разных линейных пространствах, и не должно вносить двусмысленности).

Удобство использования ковариационной функции связано, в частности, с тем, что данная функция не зависит от выбора базиса. Действительно, пусть система функций  $\tilde{\varphi}_m|_{m=1}^n$  связана с исходной соотношениями (6). Тогда ковариационная матрица для данной системы аналогично (7) будет связана с (12) следующим образом:

$$\tilde{K} = V K V^+.$$

И соответственно с учетом (7)

$$K^A = V^+ \tilde{K}^A V. \quad (15)$$

Таким образом, ковариационная функция в исходной системе может быть представлена как

$$\begin{aligned} K(x, y) &= \sum_{k,m,p,q=1}^n V_{p,k} \tilde{K}_{p,q}^A V_{q,m} \varphi_k(x) \varphi_m(y) = \\ &= \sum_{p,q=1}^n \tilde{K}_{p,q}^A \tilde{\varphi}_p(x) \tilde{\varphi}_q(y) = \tilde{K}(x, y), \end{aligned}$$

где использовалась связь (6) между различными системами. Следовательно, ковариационные функции в этих системах совпадают.

С помощью (14) выражение для дисперсии с учетом (10) и (13) может быть записано в интегральном виде

$$D\zeta_f = \iint K(x, y) f(x) \mu_0(dx) f(y) \mu_0(dy), \quad (16)$$

который также не зависит от выбора базиса.

При выводе (13) единственным ограничением на функцию являлось существование интеграла (10). В частности, если все  $\varphi_k$  имеют конечные значения в точке  $x_0$ , то можно формально положить  $f_k = \varphi_k(x_0)$ . Как можно видеть, математическое ожидание случайной величины (11) для этой функции, которую удобно обозначить как

$$\zeta(x_0) = \sum_{k,m=1}^n \varphi_k(x_0) A_{k,m}^{-1} \eta_m, \quad (17)$$

дает оценку значения  $\rho_n(x_0)$ . Аналогично (13), учитывая определение (14), можно определить дисперсию значения плотности заряда в точке:

$$D\zeta(x_0) = K(x_0, x_0). \quad (18)$$

Поскольку для любой квадратично-интегрируемой функции множество точек, в которых она неопределенна, имеет нулевую меру, оценка статистической погрешности (18) справедлива почти всюду.

Следует заметить, что оценки (11) и (17) в некотором смысле локальны и не дают представления об отклонении выборочной плотности заряда от истинного значения. В то же время хотелось бы иметь величину, которая давала бы интегральную оценку сходимости процедуры статистического моделирования. Для этого удобно рассмотреть оценку сверху всех случайных величин (11), соответствующих функциям из  $L_2(X, \mu_0)$ . Такое ограничение вполне естественно, так как все плотности лежат в этом пространстве и, следовательно, только на нем ин-

теграл (2), рассматриваемый как линейный оператор, гарантированно непрерывен. Для этих функций несложно доказать, что

$$D\zeta_f \leq \|f\|_{\mu_0}^2 \int D\zeta(x) \mu_0(dx). \quad (19)$$

Для доказательства удобно воспользоваться тем, что матрица  $K^A$  симметрична. Как указано в [6, с. 143], в этом случае существует такая унитарная матрица  $V$  (т. е.  $V^+ = V^{-1}$ ), что  $VK^AV^+$  — диагональная. Данная матрица индуцирует преобразование базиса (6). Учитывая определение (15) и независимость от базиса ковариационной функции, последнюю можно записать следующим образом:

$$K(x, y) = \sum_{m=1}^n \tilde{K}_{m,m}^A \tilde{\varphi}_m(x) \tilde{\varphi}_m(y). \quad (20)$$

Вследствие неотрицательности  $\tilde{K}$ , а следовательно,  $\tilde{K}^A$ , всегда справедливо  $K_{m,m}^A \geq 0$ . Подставляя (20) в (16) и используя неравенство Коши—Буняковского [3], можно получить следующую оценку:

$$\begin{aligned} D\zeta_f &= \sum_{m=1}^n \tilde{K}_{m,m}^A \left( \int \tilde{\varphi}_m(x) f(x) \mu_0(dx) \right)^2 \leq \\ &\leq \sum_{m=1}^n \tilde{K}_{m,m}^A \|\tilde{\varphi}_m\|_{\mu_0}^2 \|f\|_{\mu_0}^2. \end{aligned}$$

Отсюда, поскольку согласно (20)

$$\int K(x, x) \mu_0(dx) = \sum_{m=1}^n K_{m,m}^A \|\tilde{\varphi}_m\|_{\mu_0}^2,$$

получаем неравенство (19).

Впрочем, для использования в расчетах это неравенство удобнее представить в другом виде. Проинтегрировав (18), с учетом (14) и явных выражений для элементов матрицы  $A$  интеграл в неравенстве (19) можно записать в следующем виде:

$$\int K(x, x) \mu_0(dx) = \text{Sp}(A^{-1}K), \quad (21)$$

где через  $\text{Sp}$  обозначена операция вычисления следа  $\left( \text{Sp}(M) \equiv \sum_{k=1}^n M_{k,k} \right)$ . Таким образом,  $\text{Sp}(A^{-1}K)$  служит некоторой интегральной оценкой среднего квадратичного отклонения выборочного представления заряда в том смысле, что для любой квадратично-интегрируемой функции с учетом определения (9) справедливо

$$\frac{\sigma(\zeta_f)}{\|f\|_{\mu_0}} \leq \sqrt{\frac{\text{Sp}(A^{-1}K)}{N}}. \quad (22)$$

Правая часть выражения (22) не зависит от функции и имеет смысл среднего отклонения выборочной плотности заряда от предельной в пространстве  $L_2(X, \mu_0)$ . Действительно, пусть  $\rho_n^N$  — выборочная плотность, оцененная в результате  $N$  испытаний. Тогда, как показано в [1], с вероятностью (коэффициентом доверия)  $P$  выполняется соотношение

$$\left| \int (\rho_n^N(x) - \rho_n(x)) f(x) \mu_0(dx) \right| < \Phi^{-1}(P) \sigma(\zeta_f),$$

где  $\Phi$  — интеграл вероятностей [1]. Учитывая, что (22) справедливо для любой квадратично-интегрируемой функции, в том числе и для  $\rho_n^N - \rho_n$ , можно утверждать, что с коэффициентом доверия, не меньшим  $P$ , выполняется неравенство

$$\|\rho_n^N - \rho_n\|_{\mu_0} < \Phi^{-1}(P) \sqrt{\frac{\text{Sp}(A^{-1}K)}{N}}. \quad (23)$$

С учетом смысла величины (21) в дальнейшем будем называть ее *среднеквадратичной дисперсией*.

Соотношение (23) позволяет сделать оценку отклонения выборочной плотности заряда от истинной, которое, вследствие справедливости для нормы аксиомы треугольника [3], может быть представлено в следующем виде:

$$\begin{aligned} \|\rho_n^N - \rho\|_{\mu_0} &\leq \|\rho - \rho_n\|_{\mu_0} + \|\rho_n^N - \rho_n\|_{\mu_0} < \\ &< \Delta_n + \Phi^{-1}(P) \sqrt{\frac{\text{Sp}(A^{-1}K)}{N}}. \end{aligned} \quad (24)$$

В этом выражении использованы оценки отклонения (4) и (23).

Таким образом, выражения (13), (16) и (18) могут быть использованы для оценки статистической погрешности различных интегралов, а (23) и (24) — для оценки погрешности самого распределения.

На практике нередко встречается положение, когда невозможна оценка ковариационной матрицы (12). Оно связано не с какими-то принципиальными вещами, а с тем обстоятельством, что расчет этой матрицы не реализован в используемом программном продукте. Но так как оценка погрешности одновременно с расчетом основных величин является одной из сильных сторон метода статистического моделирования, хотелось бы сохранить ее и в этом случае.

Следует заметить, что существуют методы оценки погрешности без расчета дисперсии (см., например, [1, с. 90]). Применение этих методов вполне допустимо, но во многих случаях их использование менее удобно, чем реализация расчета ковариационной матрицы. Поэтому кажется наиболее приемлемым учитывать, что даже программы, которые не могут обеспечить расчет ковариационной матрицы, обычно рассчитывают ее диагональные элементы, представляющие собой дисперсии оцениваемых величин (8). Использование этих дисперсий позволяет сделать и оценки необходимых погрешностей.

Сначала удобно рассмотреть оценку погрешности в точке, которая дается выражением (18). Однако прежде необходимо заметить, что для двух неотрицательных квадратных матриц  $P$  и  $Q$  справедливо

$$\text{Sp}(PQ) \leq \text{Sp}(P) \text{Sp}(Q). \quad (25)$$

Вывод этого соотношения аналогичен выводу (19) и потому рассматриваться не будет.

Для оценки сверху дисперсии плотности в точке можно привести (18) к виду (25) таким образом, чтобы в качестве  $P$  выступала ковариационная матрица. Для этого удобно ввести обозначение

$$\psi_k(x) = \sum_{m=1}^n A_{k,m}^{-1} \varphi_m(x). \quad (26)$$

С его использованием выражение для ковариационной функции (14) для совпадающих аргументов может быть представлено в следующем виде:

$$K(x, x) = \sum_{k,m=1}^n K_{k,m} \psi_k(x) \psi_m(x) = \text{Sp}(K\Psi(x)),$$

где  $\Psi_{k,m}(x) = \psi_k(x) \psi_m(x)$ . Как можно видеть, матрица  $\Psi(x)$  неотрицательна.

Таким образом, учитывая (25), а также используя (1), (8), (10) и (12), можно привести выражение для оценки сверху дисперсии плотности в точке (26) через дисперсии оцениваемых величин:

$$D\zeta(x) \leq D_{\max}\zeta(x) = \sum_{k=1}^n D\eta(\varphi_k) \sum_{m=1}^n \psi_m^2(x). \quad (27)$$

Что касается оценки дисперсии случайной величины (11), то здесь удобно использовать мажоранту (19). При этом, если система функций  $\varphi_k$  является ортогональной, т. е. матрица  $A$

диагональна, то выражение (21) содержит только диагональные элементы ковариационной матрицы и потому вычисляется точно через дисперсии величин (8). В случае произвольной системы это неверно. Однако, учитывая то обстоятельство, что матрица  $A$  положительна, а ковариационная матрица неотрицательна, можно согласно (25) представить мажоранту следующим образом:

$$\begin{aligned} \text{Sp}(A^{-1}K) &\leq \text{Sp}(K) \text{Sp}(A^{-1}) = \\ &= \sum_{k=1}^n D\eta(\varphi_k) \sum_{m=1}^n A_{m,m}^{-1}. \end{aligned} \quad (28)$$

Впрочем, данное выражение можно получить прямым интегрированием (19) с использованием оценки сверху (27).

Следует отметить, что все оценки дисперсии, использующие ковариационную матрицу, инвариантны относительно выбора базиса на линейной оболочке функций  $\varphi_k$ , в то время как мажорирующие оценки (27) и (28), использующие только дисперсии величин (10), таким свойством не обладают.

Хочется обратить внимание на использование гистограмм. При этом в качестве функций  $\varphi_k$  рассматриваются характеристические функции областей разбиения. Для оценки погрешности используются только дисперсии соответствующих случайных величин, хотя ковариационная матрица имеет диагональный вид только в некоторых частных случаях. Однако для гистограмм пренебрежение недиагональными компонентами вполне допустимо. Например, при вычислении дисперсии плотности в точке (18) вследствие того, что для характеристических функций непесекающихся областей справедливо

$$k \neq m \Rightarrow \varphi_k(x) \varphi_m(x) = 0,$$

в ковариационной функции, представленной выражением (14), остаются только диагональные компоненты ковариационной матрицы. В случае же необходимости оценки погрешности интеграла некоторой функции вполне допустимо для этого использовать неравенство (19), которое с учетом (21), а также диагонального вида матрицы  $A$  — следствия ортогональности характеристических функций — содержит только диагональные компоненты ковариационной матрицы.

Представляется полезным сделать оценку удобства использования гистограмм для представления гладких функций. Для этого следу-

ет заметить, что при достаточно общих предположениях при большой размерности пространства представления ( $n$ ) среднеквадратичная дисперсия пропорциональна этой размерности:  $\text{Sp}(A^{-1}K) \sim n$ .

Для определенности можно рассмотреть одномерное пространство параметров с лебеговской базовой мерой. В этом случае, если плотность заряда является достаточно гладкой, например аналитической, функцией, можно показать, что погрешность представления (4) гистограммой обратно пропорциональна  $n$ :

$$\Delta_n^\Gamma \sim \frac{1}{n}.$$

В то же время, как показано в [7, с. 149], для погрешности представления (4) степенными функциями существует следующая оценка сверху:

$$\Delta_n^c \leq \frac{C_p}{n^p},$$

где  $p$  — произвольное натуральное число, а  $C_p$  — связанная с ним константа (в общем случае ограничение  $p$  связано с существованием непрерывных производных этой степени).

Чтобы оценить выигрыш в объеме вычислений, который получается при использовании степенных функций вместо гистограммы, в качестве критерия удобно использовать равенство среднеквадратичных отклонений выборочной плотности. Учитывая (23) и считая, что объемы вычислений пропорциональны объемам выборки, их отношение можно получить в следующем виде:

$$\frac{N_\Gamma}{N_c} = \frac{\text{Sp}(A_\Gamma^{-1}K_\Gamma)}{\text{Sp}(A_c^{-1}K_c)} \geq C\Delta^{1/p-1}, \quad (29)$$

где индексом "с" обозначены величины, связанные со степенным представлением плотности, а индексом "Г" — с гистограммным. Так как неизвестно значение константы  $C$ , сложно сказать, чему будет равно это отношение в конкретном случае. Однако можно утверждать наверняка, что при малом значении погрешности представления объем вычислений, необходимый при использовании гистограммы, будет намного больше аналогичного объема при использовании степенных функций.

Представляется необходимым привести пример, показывающий возможность использования различных оценок дисперсии для системы непрерывных функций.

Пусть на интервале  $x \in (-1, 1)$  задана функция

$$\rho(x) = \frac{3}{10}(2 + x - x^2), \quad (30)$$

являющаяся, как можно видеть, плотностью распределения. В качестве случайной величины (1) для произвольной функции  $f$  рассматривается

$$\eta(f) = f(x), \quad (31)$$

где  $x$  — случайное значение, разыгранное в соответствии с плотностью  $\rho(x)$ . Пусть в качестве базовой меры используется мера Лебега на интервале  $(-1, 1)$ :

$$\int f(x) \mu_0(dx) = \int_{-1}^1 f(x) dx.$$

Как можно видеть, плотность точно представляется многочленом степени не ниже двух. С учетом этого обстоятельства будут рассматриваться следующие системы функций:  $\varphi_{1,k}(x) = x^{k-1}$ ;  $\varphi_{2,k}(x) = P_{k-1}(x)$ , где  $P_k(x)$  — многочлены Лежандра, ортогональные на интервале  $(-1, 1)$ . Следует заметить, что несимметричный относительно отражения от оси ординат вид плотности распределения выбран для того, чтобы исключить влияние симметрии базисных функций. Данные системы с ограниченным значением  $k$  представляют собой различные базисы в пространстве многочленов степени ниже максимального значения  $k$ . Для этих систем, в частности, с учетом данной вероятностной модели, определяемой выражениями (30) и (31), несложно получить явное выражение для дисперсии плотности в точке (18) и ее мажоранты (27), а также их интегралы — (21) и (28) соответственно. При этом, как было указано выше, выражения (21) в обеих системах одинаковы, а (28) различны.

Для выявления характерных особенностей при использовании разных оценок и разных систем функций удобно рассмотреть таблицу, в которой приведены интегральные оценки дисперсий для разных максимальных степеней пространства многочленов (*размерность пространства = степень + 1*). Здесь через  $K_m$ ,  $A_m$  обозначены матрицы, соответствующие системам функций  $\varphi_{m,k}|_{k=1}^n$  (так как во втором столбце приводится величина, инвариантная относительно замены базиса, то индекс в этом случае опущен). Для этих систем функций значение погрешности представления (4)  $\Delta_n = 0$  и ошибка определяется только статистическим разбросом. Как

### Оценки среднеквадратичной дисперсии

Степень	$\text{Sp}(A^{-1}K)$	$\frac{\text{Sp}(K_1) \times}{\times \text{Sp}(A_1^{-1})}$	$\frac{\text{Sp}(K_2) \times}{\times \text{Sp}(A_2^{-1})}$
2	0,79	2,6	1,8
3	1,2	15	4,3
5	2,1	501	13
9	3,9	$5,4 \times 10^5$	47

можно видеть, интегральное значение дисперсии (19), представленное во втором столбце таблицы, растет приблизительно по линейному закону с ростом степени многочлена. Однако рост мажорантной дисперсии, представленной в третьем и четвертом столбцах, происходит значительно быстрее. Впрочем, для ортогональной системы с учетом того, что согласно (9) ошибка определяется квадратным корнем из дисперсии, этот рост вполне приемлем. Таким образом, можно предположить, что по крайней мере в случаях приближения плотности заряда степенными функциями использование мажорантной оценки дисперсии может быть допустимо только при небольших размерностях пространства представления либо при использовании системы ортогональных многочленов в качестве базиса.

Также представляется полезным оценить выигрыш в объеме вычислений, который получается для гладких распределений при использовании степенных функций вместо гистограммы. Для сравнения необходимо задаться точностью представления. В качестве этой величины удобно использовать безразмерное отношение  $\Delta_n / \|\rho\|$ . Не приводя явно громоздких вычислений (которые, впрочем, несложно повторить), можно привести следующие значения минимального числа интервалов гистограммы и дисперсии данного представления:

$$\frac{\Delta_n}{\|\rho\|} < 0,01 \Rightarrow n = 50 \Rightarrow \text{Sp}(A_\Gamma^{-1}K_\Gamma) \approx 25;$$

$$\frac{\Delta_n}{\|\rho\|} < 0,001 \Rightarrow n = 494 \Rightarrow \text{Sp}(A_\Gamma^{-1}K_\Gamma) \approx 247.$$

Если взять максимальное и минимальное значения дисперсии из второго столбца таблицы для знаменателя в формуле (29), то отношение объемов вычислений лежит в следующих интервалах значений:

$$\frac{\Delta_n}{\|\rho\|} < 0,01 \Rightarrow \frac{N_\Gamma}{N_c} \approx 6 \div 32;$$

$$\frac{\Delta_n}{\|\rho\|} < 0,001 \Rightarrow \frac{N_\Gamma}{N_c} \approx 63 \div 313.$$

Таким образом, в данной задаче использовать гистограммное представление разумно в случае достаточно грубого приближения, порядка нескольких процентов. Если же относительная погрешность должна быть ниже, то использование такого представления приводит к значительному росту вычислительных затрат.

### Заключение

Как было отмечено, при расчете основной задачи, т. е. при оценке функционалов (8), никак не используется базовая мера. Поэтому на этапе обработки рассчитанных величин появляется возможность выбрать эту меру таким образом, чтобы представление плотности заряда было в каком-то смысле наилучшим. Понятие наилучшего представления можно ввести разным образом, однако оно должно включать то условие, что погрешность  $\Delta_n$ , определяемая выражением (4), будет в каком-то смысле минимальной на некотором множестве допустимых базовых мер. Впрочем, задача наилучшего представления, с одной стороны, в общем случае достаточно обширна, а с другой, в большинстве частных случаев очевидна. Поэтому ее рассмотрение выходит за рамки настоящей работы.

В заключение представляется важным привести основные выражения, используемые при практическом решении задачи приближения. Так, само приближение (4) получается при решении системы уравнений (5). Дисперсия для произвольной функции дается выражениями (13) или (16). Дисперсия плотности заряда в точке определяется выражением (18), а среднеквадратичная дисперсия — (21). Оценки сверху даются выражениями (27) для дисперсии в точке и (28) — для среднеквадратичной дисперсии. При этом последние две величины в отличие от всех предыдущих зависят от выбора базиса в пространстве представления. Здесь следует иметь в виду, что оценка погрешности с помощью дисперсии в точке имеет значение, если плотность непрерывна, а  $\Delta_n$  мала по сравнению со среднеквадратичной дисперсией.

Следует отметить, что используемые в настоящей работе оценки позволяют проводить численные сравнения погрешностей разных приближений, что, в свою очередь, позволяет выбирать из этих приближений наилучшее.

Кроме того, хотя некоторый упор делался на использование гладких распределений, описанная методика не в меньшей мере пригодна и для

представления зарядов, сингулярных относительно меры Лебега. Необходимо только помнить, что базовая мера должна удовлетворять требованиям, описанным в начале статьи.

Также хотелось бы еще раз обратить внимание, что использование гладких функций для представления гладких распределений при условии большой точности, согласно (29), сильно снижает необходимый объем вычислений по сравнению с кусочно-постоянным представлением. Думается, что при повышении требования к точности важность практического использования гладких оценок будет возрастать.

### Список литературы

1. *Соболев И. М.* Численные методы Монте-Карло. М.: Наука, 1973.
2. *Ермаков С. М.* Метод Монте-Карло и смежные вопросы. М.: Наука, 1971.
3. *Колмогоров А. Н., Фомин С. В.* Элементы теории функций и функционального анализа. М.: Наука, 1976.
4. *Фролов А. С., Ченцов Н. Н.* О вычислении методом Монте-Карло определенных интегралов, зависящих от параметра // Журнал вычисл. мат. и мат. физ. 1962. Т. 2, № 4. С. 714—717.
5. *Сизова А. Ф.* Оценка плотности делений в сферическом реакторе // Метод Монте-Карло в проблеме переноса излучений. Атомиздат, 1967.
6. *Гельфанд И. М.* Лекции по линейной алгебре. М.: Наука, 1971.
7. *Суетин П. К.* Классические ортогональные многочлены. М.: Физматлит, 2005.

Статья поступила в редакцию 27.11.08.

---