Сибирский математический журнал
Март—апрель, 2004. Том 45, № 2

УДК 519.245

ОПТИМИЗАЦИЯ ВЕСОВЫХ МЕТОДОВ МОНТЕ–КАРЛО ПО ВСПОМОГАТЕЛЬНЫМ ПЕРЕМЕННЫМ

Г. А. Михайлов, И. Н. Медведев

Аннотация: Рассматриваются задачи, математическая модель которых определяется некоторой, обрывающейся с вероятностью 1, цепью Маркова, причем необходимо оценивать линейные функционалы от решения интегрального уравнения 2-го рода с соответствующими субстохастическим ядром и свободным элементом [1]. Для построения весовых модификаций численного статистического моделирования в число координат фазового пространства включаются вспомогательные переменные, случайные значения которых функционально определяют переходы в исходной цепи. После реализации каждой вспомогательной случайной величины вес домножается на отношение соответствующих плотностей исходного и численно моделируемого распределения. Решается задача минимизации дисперсий оценок линейных функционалов путем подбора моделируемого распределения первой по порядку вспомогательной случайной величины.

Ключевые слова: метод Монте-Карло, субстохастическое ядро, обобщенная плотность, цепь Маркова, интегральное уравнение, линейный функционал, вспомогательные веса, оптимизация моделирования, зкспоненциальное преобразование.

1. Вводная информация

1.1. Математическая модель ряда прикладных задач строится на основе рассмотрения некоторого скачкообразного обрывающегося с вероятностью единица однородного марковского процесса (см., например, [1]). При этом траектория процесса вполне определяется ее состояниями в моменты скачков, т. е. фактически можно рассматривать обрывающуюся однородную цепь Маркова с заданной переходной функцией $P(x, S)$, где $x \in X$, X — m-мерное евклидово пространство, $S \subset X$ — измеримое по Лебегу множество. Для построения весовых алгоритмов моделирования целесообразно рассматривать соответствующую условной мере $P(x', S)$ обобщенную субстохастическую плотность перехода $k(x', x)$. Обобщенная плотность распределения $k(x', x)$ определяется для любого $x' \in X$ равенством

$$\int\limits_X h(x)P(x', dx) = \int\limits_X k(x', x)h(x)\,dx \quad \forall h \in C_0(X),$$

где $C_0(X)$ — множество непрерывных ограниченных функций. Отметим, что здесь и далее рассматриваются и ненормированные (т. е. не обязательно вероятностные) распределения.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант 03-01-00040), интеграционного гранта СО РАН–2003–№2, а также программы «Ведущие научные школы» (грант НШ–1271.2003.1).

Настоящая работа ориентирована на приложения в теории переноса частиц, в которой кроме измеримых плотностей, соответствующих абсолютно непрерывным распределениям, используются также «дельта-функции», означающие интегрирование по некоторым многообразиям меньшей сравнительно с m размерности (см., например, [1, разд. 2.1.1). Использовать обобщенные плотности (вместо интегрирования по соответствующим мерам) в теории статистического моделирования предложил Н. Н. Ченцов [2] в связи с тем, что такой подход упрощает построение и реализацию модификаций моделирования. Это важно и для настоящей работы, так как в ней рассматриваются дополнительные фазовые переменные, причем базовые переменные — координаты и скорости — связаны с дополнительными функционально, т. е. их условные распределения определяются соответствующими «дельта-функциями».

Предполагается, что

$$\int k(x', x)\, dx = q(x') \leq 1 - \delta < 1.$$

Величина $q(x')$ имеет смысл вероятности необрыва траектории в заданной точке x'. Вследствие последнего неравенства цепь обрывается с вероятностью единица и среднее число состояний конечно.

Итак, рассматривается однородная обрывающаяся цепь Маркова

$$x_0, x_1, \dots, x_N,$$

определяемая плотностью $f(x)$ распределения начального состояния x_0 и субстохастической обобщенной плотностью перехода $k(x', x)$. Здесь N — номер состояния, в котором реализуется обрыв траектории (иначе, *момент остановки*). Ясно, что обобщенная плотность распределения состояний, непосредственно следующих за начальным, выражается равенством

$$\varphi_1(x) = \int f(x')k(x', x)\, dx' = [Kf](x).$$

Следовательно, обобщенная плотность распределения фазовых состояний цепи

$$\varphi(x) = \sum_{n=0}^{\infty} \varphi_n(x),$$

где $\varphi_n(x)$ — плотность распределения состояний номера n, представляет собой ряд Неймана для следующего интегрального уравнения 2-го рода:

$$\varphi(x) = \int\limits_X k(x', x)\varphi(x')\, dx' + f(x) \tag{1.1}$$

или $\varphi = K\varphi + f$. Это уравнение можно рассматривать в пространстве $N_1(X)$ обобщенных плотностей мер ограниченной вариации, так как ряд $\sum\limits_{n=0}^{\infty}(K^n f, h)$ в силу условия $q(x') \leq 1 - \delta$ сходится для любой $h \in C_0(X)$. Для заданной функции $h \in C_0(X)$ рассмотрим также сопряженное уравнение $\varphi^* = K^*\varphi^* + h$, где

$$[K^*\varphi^*](x') = \int\limits_X k(x', x)\varphi^*(x)\, dx.$$

Предполагается, что $K^* \in [C_0(X) \to C_0(X)]$.

Справедливо представление

$$\varphi^* = \sum_{n=0}^{\infty} K^{*n} h = \sum_{n=0}^{\infty} (K^n \delta, h),$$

где $\delta(\cdot)$ — обобщенная плотность локализованного в соответствующей точке источника.

Методы Монте-Карло обычно используются для оценки линейных функционалов вида

$$I_h = (\varphi, h) = \int\limits_X \varphi(x) h(x)\, dx = \sum_{n=0}^{\infty} (K^n f, h), \quad h \in C_0(X).$$

Если реализуется прямое моделирование исходной цепи Маркова, то для оценки величины I_h используется соотношение

$$I_h = \mathbf{E}\left[\sum_{n=0}^{N} h(x_n)\right].$$

Однако можно моделировать и другую, вспомогательную, цепь Маркова с плотностью перехода $p(x', x)$, взаимно регулярной с $k(x', x)$, и начальной плотностью $\pi(x)$, взаимно регулярной с $f(x)$. Кроме того, предполагается, что $p(x', x) \neq 0$ и $\pi(x) \neq 0$ на носителях функций $k(x', x)$ и $f(x)$ соответственно. При выполнении этих «общих условий несмещенности» отношения $k(x', x)/p(x', x)$ и $f(x)/\pi(x)$ имеют смысл и определяются отношениями измеримых сомножителей рассматриваемых функций. Это позволяет ввести вспомогательные веса по формулам

$$Q_0(x_0) = \frac{f(x_0)}{\pi(x_0)}, \quad Q_n = Q_{n-1} \frac{k(x_{n-1}, x_n)}{p(x_{n-1}, x_n)}. \tag{1.2}$$

Полагая

$$\xi = \sum_{n=0}^{N} Q_n h(x_n),$$

имеем $I_h = \mathbf{E}\xi$ (см., например, [1]). Величина ξ представляет собой весовую «оценку по столкновениям». Введем величину $\rho(K)$ по формуле

$$\rho(K) = \rho(K^*) = \inf_n \|K^{*n}\|^{1/n}.$$

Обозначим символом K_p оператор с ядром $k^2(x', x)/p(x', x)$. Известно [1], что если $\rho(K_p) < 1$ и $f^2/\pi \in N_1(X)$, то $\mathbf{D}\xi < \infty$. Известно также, что если $h(x) \geq 0$ и

$$p(x', x) = \frac{k(x', x)\varphi^*(x)}{[K^*\varphi^*](x')}, \quad \pi(x) = \frac{f(x)\varphi^*(x)}{(f, \varphi^*)},$$

то $\mathbf{D}\xi = 0$. Если вместо $\varphi^*(x)$ здесь использовать функцию

$$g(x) = C\varphi^*(x)[1 + \varepsilon(x)], \quad |\varepsilon(x)| \leq \varepsilon,$$

то при малом ε величина $\mathbf{D}\xi$ мала [3]. Соответствующие алгоритмы объединяются под названием «моделирование по ценности» [1, 3]. Функцию $\varphi^*(\cdot)$ в теории весовых методов Монте-Карло принято называть «функцией ценности» в связи с ее вероятностным представлением:

$$\varphi^*(x) = h(x) + \mathbf{E} \sum_{n=1}^{N} Q_n h(x_n) \quad \text{при } x_0 \equiv x. \tag{1.3}$$

Допустимы кусочно-непрерывные $h(\cdot)$, если интегралы $(K^n f, h)$ определены.

1.2. Как правило, переход $x' \to x$ осуществляется в результате выбора совокупности значений вспомогательных случайных величин (может быть, векторных): $\mathbf{t} = (t_1, \dots, t_m) \in T$, т. е. справедливо представление

$$k(x', x) = \int_{\tau} k_1(x', t_1) k_2((x', t_1), t_2) \dots k_m((x', t_1, \dots, t_{m-1}), t_m) \delta(x - x(x', \mathbf{t}))\, d\mathbf{t}.$$

Здесь $x(x', \mathbf{t})$ — функция, определяющая новое значение стандартных евклидовых координат через x' и значения вспомогательных переменных $\mathbf{t}$.

Обычно построение весов непосредственно по формулам (1.2) для решения достаточно сложных задач оказывается практически невозможным. Для преодоления этого затруднения целесообразно перейти к модифицированному фазовому пространству $T \times X$, точками которого являются совокупности $(\mathbf{t}, x)$. Соответствующее субстохастическое ядро имеет вид

$$\mathbf{k}((\mathbf{t}', x'), (\mathbf{t}, x)) = \delta(x - x(x', \mathbf{t})) \prod_{i=1}^{m} k_i((x', t_1, \dots, t_{i-1}), t_i), \tag{1.4}$$

причем (x', t_1, t_0) следует рассматривать как x'. Введем обозначения: $\mathbf{K}$ — интегральный оператор с ядром (1.4); $\mathbf{f}(\mathbf{t}, x) = f_0(\mathbf{t}) \cdot f(x)$, где $f_0(\mathbf{t})$ — некоторая плотность вероятностей в T;

$$\varphi_t = \mathbf{K}\varphi_t + \mathbf{f}; \quad \mathbf{I}_h = (\varphi_t, h); \quad \varphi_t^* = \mathbf{K}^* \varphi_t^* + h.$$

Справедливы следующие равенства:

$$\varphi(x) = \int_T \varphi_t(\mathbf{t}, x) d\mathbf{t}, \quad \mathbf{I}_h = I_h, \quad \varphi_t^*(\mathbf{t}, x) \equiv \varphi^*(x).$$

Первое равенство получается почленным интегрированием модифицированного интегрального уравнения по $\mathbf{t}$, второе — частичным интегрированием по $\mathbf{t}$ в интеграле (φ_t, h), а третье вытекает из вероятностного представления функции ценности (1.3).

Для построения модифицированной оценки по столкновениям ξ_t определяются вспомогательная переходная плотность $\mathbf{p}(\cdot, \cdot)$ вида (1.4), начальная плотность $\pi_t(\mathbf{t}, x) = f_0(\mathbf{t})\pi(x)$ и соответствующие веса $\{\mathbf{Q}_n\}$, которые получаются домножением на весовые множители после каждого элементарного перехода. При выполнении указанных в п. 1.1 общих условий несмещенности (для каждого элементарного перехода) имеем

$$I_h = \mathbf{E}\xi_t, \quad \text{где } \xi_t = \sum_{n=0}^{N} \mathbf{Q}_n h(x_n),$$

причем $\mathbf{D}\xi_t < +\infty$, если $\rho(\mathbf{K}_p) < 1$, где $\mathbf{K}_p$ — интегральный оператор с ядром $\mathbf{k}^2(\cdot, \cdot)/\mathbf{p}(\cdot, \cdot)$.

Введем для $i = 1, \dots, m$ вспомогательные функции ценности

$$\varphi_i^*(x', t_1, \dots, t_i) = \underbrace{\int \dots \int}_{m-i+1} \delta(x - x(x', \mathbf{t})) \left[\prod_{j=i+1}^{m} k_j((x', t_1, \dots, t_{j-1}), t_j) \right] \times \varphi^*(x)\, dt_{i+1} \dots dt_m dx, \tag{1.5}$$

полагая $\prod\limits_{m+1}^{m} \equiv 1$. Нетрудно заметить, что $\varphi_i^*(x', t_1, \dots, t_{i-1}, t_i)$ представляет собой условное математическое ожидание величины ξ_t при условии, что траектория цепи начинается во вспомогательной точке $(x', t_1, \dots, t_i)$.

Теорема 1.1. *Пусть* $h(\cdot) \geq 0$. *Если для* $i = 2, \ldots, m$

$$p_i((x', t_1 \ldots, t_{i-1}), t_i) = \frac{k_i((x', t_1, \ldots, t_{i-1}), t_i)\varphi_i^*(x', t_1, \ldots, t_{i-1}, t_i)}{\varphi_{i-1}^*(x', t_1, \ldots, t_{i-2}, t_{i-1})},$$
$$p_1(x', t_1) = \frac{k_1(x', t_1)\varphi_1^*(x', t_1)}{\varphi^*(x') - h(x')} \quad \text{и} \quad \pi(x) = \frac{f(x)\varphi^*(x)}{I_h}, \tag{1.6}$$

то $\mathbf{D}\xi_t = 0$.

Доказательство. Нормированность условных плотностей $p_i(\cdot, \cdot)$ здесь следует непосредственно из определения функций ценности $\varphi_i^*(\cdot)$. Подставляя указанные плотности в выражение для $\mathbf{p}(\cdot, \cdot)$, убеждаемся, что теорема является частным случаем приведенного в п. 1.1 общего утверждения об оценке по столкновениям с нулевой дисперсией.

Отметим, что, меняя порядок выбора величин $t_1, \ldots, t_m$, т. е. способ факторизации субстохастического ядра, можно строить различные интегральные уравнения и на их основе весовые оценки заданного функционала I_h. В частности, можно осуществлять сдвиг вдоль цепочки элементарных переходов, т. е. фиксировать фазовое состояние после выбора t_i, $i < m$. Иногда это может упростить выражение вспомогательного веса и параметрический анализ результатов.

Как будет видно из дальнейшего, использование «ценностных» плотностей вида (1.4) лишь для части вспомогательных переменных может быть малоэффективным.

2. Оптимизация моделирования по части переменных

Здесь будут использованы две, вообще говоря, векторные вспомогательные случайные величины t_1, t_2, т. е. ядро имеет вид

$$\mathbf{k}((\mathbf{t}, x), (\mathbf{t}', x')) = \delta(x' - x'(x, \mathbf{t}'))k_1(x, t_1')k_2((x, t_1'), t_2').$$

В дальнейшем вспомогательные переменные будут рассматриваться всегда, поэтому будут использоваться более простые обозначения, принятые в теории весовых оценок (см.[1]): $\xi, \varphi, \varphi^*, I_h, K, K_p$ и т. д. Выполняются соотношения

$$\xi = \frac{f(x_0)}{\pi(x_0)}\xi_{x_0},\ \xi_x = h(x) + \sum_{n=1}^{N} Q_n h(x_n), \quad f \in N_1(X),\ h \in C_b(X) = C_0(X) \cap C_+(X),$$

$$\mathbf{E}\xi_x = \varphi^*(x), \quad \mathbf{E}\xi^2 = \int\limits_X \frac{f^2(x)}{\pi(x)} \mathbf{E}{\xi_x}^2\, dx.$$

Поэтому целесообразно рассмотреть задачу равномерной относительно $x \in \mathbf{X}$ минимизации величины $\mathbf{E}\xi_x^2$. Согласно теореме 1.1 использование «ценностных» плотностей вида (1.6) для всех элементарных переходов дает абсолютный минимум: $\mathbf{E}\xi_x^2 = (\varphi^*(x))^2\ \forall x \in X$. Практически такая глобальная оптимизация моделирования весьма затруднительна, поэтому важно рассмотреть возможность уменьшения величины $\mathbf{E}\xi_x^2$ путем оптимального подбора плотности распределения части вспомогательных случайных величин, например t_1.

Рассмотрим цепь Маркова с субстохастической плотностью перехода:

$$\mathbf{p}((\mathbf{t}, x), (\mathbf{t}', x')) = \delta(x' - x'(x, \mathbf{t}'))p_1(x, t_1')p_2((x, t_1'), t_2').$$

Пусть дополнительно

$$\int\limits_{T_1} k_1(x,t'_1)dt'_1 \equiv \int\limits_{T_1} p_1(x,t'_1)dt'_1 \equiv 1, \quad \int\limits_{T_2} \frac{k_2^2((x,t'_1),t'_2)}{p_2((x,t'_1),t'_2)}\, dt'_2 \le q < 1. \qquad (2.1)$$

Функция $\mathbf{E}\xi_x^2$ выражается рядом Неймана для уравнения [3]

$$\chi = h(2\varphi^* - h) + K_p\chi.$$

Поскольку выполнено соотношение

$$\int\limits_X \delta(x' - x'(x,\mathbf{t}'))\mathbf{E}\xi_{x'}^2\, dx' = \mathbf{E}\xi_{x'(x,\mathbf{t}')}^2 = \mathbf{E}\xi_{x'}^2,$$

то

$$\mathbf{E}{\xi_x}^2 = h(x)[2\varphi^*(x) - h(x)] + \int\limits_{T_1} \frac{k_1^2(x,t'_1)}{p_1(x,t'_1))}\left[\int\limits_{T_2} \frac{k_2^2((x,t'_1),t'_2)}{p_2((x,t'_1),t'_2)}\mathbf{E}\xi_{x'}^2\, dt'_2\right] dt'_1. \qquad (2.2)$$

Рассмотрим следующую задачу: *для фиксированной плотности* $p_2((x,t'_1),t'_2)$ *определить плотность* $p_1(x,t'_1)$ *таким образом, чтобы функция* $\mathbf{E}\xi_x^2$ *для любого* $x \in X$ *была минимальной при выполнении общих условий несмещенности.* Такую плотность $p_1(x,t'_1)$ будем называть *равномерно наилучшей.*

Согласно принципу выборки по важности [1, 3] интеграл в правой части (2.2) минимален и равен

$$\left\{\int\limits_{T_1} k_1(x,t'_1)\left[\int\limits_{T_2} \frac{k_2^2((x,t'_1),t'_2)}{p_2((x,t'_1),t'_2)}\mathbf{E}\xi_{x'}^2\, dt'_2\right]^{1/2} dt'_1\right\}^2$$

при

$$p_1(x,t'_1) = \frac{k_1(x,t'_1)\left[\int\limits_{T_2} \frac{k_2^2((x,t'_1),t'_2)}{p_2((x,t'_1),t'_2)}\mathbf{E}\xi_{x'}^2\, dt'_2\right]^{1/2}}{\int\limits_{T_1} k_1(x,t'_1)\left[\int\limits_{T_2} \frac{k_2^2((x,t'_1),t'_2)}{p_2((x,t'_1),t'_2)}\mathbf{E}\xi_{x'}^2\, dt'_2\right]^{1/2} dt'_1}.$$

Таким образом,

$$g(x) = h(x)[2\varphi^*(x) - h(x)] + \left\{\int\limits_{T_1} k_1(x,t'_1)\left[\int\limits_{T_2} \frac{k_2^2((x,t'_1),t'_2)}{p_2((x,t'_1),t'_2)}\mathbf{E}\xi_{x'}^2\, dt'_2\right]^{1/2} dt'_1\right\}^2 \le g_0(x) = \mathbf{E}\xi_x^2. \qquad (2.3)$$

Заметим, что общие условия несмещенности не нарушаются, так как

$$\mathbf{E}\xi_x^2 \ge (\mathbf{E}\xi_x)^2 = (\varphi^*(x))^2.$$

В операторном виде соотношение (2.3) запишем так: $g_1 = G(g_0) \le g_0$. Подставив сюда g_1 вместо g_0, получим

$$g_2 = G(G(g_0)) = G(g_1) \le g_0.$$

Заметим, что такая подстановка имеет смысл оптимизации плотности распределения второго перехода $x' \to x''$ с вытекающей отсюда повторной оптимизацией первого перехода.

Далее, имеем

$$g_n(x) = [G^n(g_0)](x) \downarrow g(x)\ \forall x,$$

причем $g(x) \ge (\varphi^*(x))^2$ и

$$g(x) = h(x)[2\varphi^*(x) - h(x)] + \left\{\int\limits_{T_1} k_1(x,t'_1)\left[\int\limits_{T_2} \frac{k_2^2((x,t'_1),t'_2)}{p_2((x,t'_1),t'_2)} g(x')dt'_2\right]^{1/2} dt'_1\right\}^2. \qquad (2.4)$$

Лемма 2.1. *Если $g \in C_b(X)$ — решение уравнения (2.4), то для цепи Маркова с*

$$p_1(x,t_1') = \frac{k_1(x,t_1')\Big[\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')} g(x')\,dt_2'\Big]^{1/2}}{\int\limits_{T_1} k_1(x,t_1')\Big[\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')} g(x')\,dt_2'\Big]^{1/2} dt_1'} \qquad (2.5)$$

имеем $\mathbf{E}\xi_x^2 = g(x)$.

Доказательство. Обозначим через K_g интегральный оператор (2.2) при выполнении (2.5), т. е.

$$[K_g f](x) = \int\limits_{T_1} \bigg\{ \int\limits_{T_1} k_1(x,t_1') \bigg[\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')} g(x')\,dt_2' \bigg]^{1/2} dt_1' \bigg\}$$

$$\times \int\limits_{T_2} k_1(x,t_1') \frac{k_2^2((x,t_1'),t_2') f(x')}{p_2((x,t_1'),t_2')\Big[\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')} g(x') dt_2'\Big]^{1/2}}\, dt_2' dt_1'.$$

Очевидно, из (2.4) следует равенство $g = h(2\varphi - h) + K_g g$. При построении итераций этого уравнения начиная с $g_0 = g$ получаем

$$g = \sum_{n=0}^{\infty} K_g^n [h(2\varphi^* - h)] + \lim_{n\to\infty} K_g^n g.$$

Рассмотрим функцию $K_g^2 g$. По определению

$$[K_g^2 g](x) = \int\limits_{T_1}\int\limits_{T_1} \bigg\{ \int\limits_{T_1} k_1(x,t_1') \bigg[\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')} g(x') dt_2' \bigg]^{1/2} dt_1' \bigg\}$$

$$\times \int\limits_{T_2} k_1(x,t_1') \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')\Big[\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')} g(x') dt_2'\Big]^{1/2}}$$

$$\times \bigg\{ \int\limits_{T_1} k_1(x',t_1'') \bigg[\int\limits_{T_2} \frac{k_2^2((x',t_1''),t_2'')}{p_2((x',t_1''),t_2'')} g(x'') dt_2'' \bigg]^{1/2} dt_1'' \bigg\}$$

$$\times \int\limits_{T_2} k_1(x',t_1'') \frac{k_2^2((x',t_1''),t_2'') g(x'')}{p_2((x',t_1''),t_2')\Big[\int\limits_{T_2} \frac{k_2^2((x',t_1''),t_2'')}{p_2((x',t_1''),t_2'')} g(x'')\,dt_2''\Big]^{1/2}}\, dt_2'' dt_1'' dt_2' dt_1'.$$

Из уравнения (2.4) вытекает, что

$$\bigg\{ \int\limits_{T_1} k_1(x',t_1'') \bigg[\int\limits_{T_2} \frac{k_2^2((x',t_1''),t_2'')}{p_2((x',t_1''),t_2'')} g(x'') dt_2'' \bigg]^{1/2} dt_1'' \bigg\} \le g^{1/2}(x').$$

Следовательно,

$$\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2') \Big\{ \int\limits_{T_1} k_1(x',t_1'') \Big[\int\limits_{T_2} \frac{k_2^2((x',t_1''),t_2'')}{p_2((x',t_1''),t_2'')} g(x'')\,dt_2'' \Big]^{1/2} dt_1'' \Big\}}{p_2((x,t_1'),t_2')\Big[\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')} g(x')\,dt_2'\Big]^{1/2}}\, dt_2'$$

$$\le \int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2') g^{1/2}(x')}{p_2((x,t_1'),t_2')\Big[\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')} g(x')\,dt_2'\Big]^{1/2}}\, dt_2' \le q^{1/2},$$

так как
$$\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')}dt_2' = \tilde q(x,t_1') \le q$$
и
$$\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')}g^{1/2}(x')dt_2' = \tilde q(x,t_1')\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{\tilde q(x,t_1')p_2((x,t_1'),t_2')}g^{1/2}(x')\,dt_2'$$
$$\equiv \tilde q(x,t_1')\mathbf{E}'g^{1/2} \le \tilde q(x,t_1')(\mathbf{E}'g)^{1/2} = \tilde q^{1/2}(x,t_1')\Bigg[\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')}g(x')\,dt_2'\Bigg]^{1/2}$$
$$\le q^{1/2}\Bigg[\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')}g(x')\,dt_2'\Bigg]^{1/2}.$$

Кроме того,
$$\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')}g(x')dt_2' \le q\|g\|.$$
Поэтому $K_g^2 g \le q^{1/2}q\|g\|$. Нетрудно аналогично получить неравенство $K_g^n g \le q^{(n-1)/2}q\|g\|$. Таким образом, функция g равна ряду Неймана для уравнения (2.2) при выполнении (2.5). Лемма 2.1 доказана

Теорема 2.1. *Пусть $h \in C_b(x)$ и выполнены соотношения* (2.1). *Тогда решение g^* уравнения* (2.4) *в $C_b(X)$ существует и единственно.*

Доказательство. Уравнение (2.4) можно записать в виде
$$g(x) = h(x)[2\varphi^*(x) - h(x)] + (K_g g)^2$$
или
$$u = \sqrt{h(x)[2\varphi^*(x) - h(x)] + (Ku^2)^2} = \widetilde G(u).$$
Здесь $u = g^{1/2}$, а K — интегральный оператор вида
$$[Kf](x) = \int\limits_{T_1} k_1(x,t_1')\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')f(x')}{p_2((x,t_1'),t_2')\Big[\int\limits_{T_2}\frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')}g(x')\,dt_2'\Big]^{1/2}}dt_2'\,dt_1'.$$

Заметим, что
$$Ku^2(x) = \int\limits_{T_1} k_1(x,t_1')\Bigg[\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')}u^2(x')dt_2'\Bigg]^{1/2}dt_1' = \int\limits_{T_1} k_1(x,t_1')\|\tilde u\|_{L_2(X)}dt_1',$$
где
$$\tilde u \equiv u(x')k_2((x,t_1'),t_2')/\sqrt{p_2((x,t_1'),t_2)}.$$
Поэтому, учитывая неотрицательность функции $h(2\varphi^* - h)$, получаем
$$|\widetilde G(\tilde u_1) - \widetilde G(\tilde u_2)| \le \left|K\tilde u_1^2 - K\tilde u_2^2\right|$$
$$= \left|\int\limits_{T_1} k_1(x,t_1')\|\tilde u_1\|_{L_2(X)}dt_1' - \int\limits_{T_1} k_1(x,t_1')\|\tilde u_2\|_{L_2(X)}dt_1'\right|$$
$$\le \left|\int\limits_{T_1} k_1(x,t_1')|\|\tilde u_1\|_{L_2(X)} - \|\tilde u_2\|_{L_2(X)}|\,dt_1'\right|$$
$$\le \int\limits_{T_1} k_1(x,t_1')\Bigg[\int\limits_{T_2} \frac{k_2^2((x,t_1'),t_2')}{p_2((x,t_1'),t_2')}(u_1(x') - u_2(x'))^2\,dt_2'\Bigg]^{1/2}dt_1' \le q^{1/2}\|u_1 - u_2\|.$$

Таким образом, $\widetilde{G}$ — оператор сжатия в $C_b(X)$, что обеспечивает существование и единственность решения g^*. Теорема 2.1 доказана.

Заметим, что плотность (2.5) дает равномерно наименьшее значение $\mathbf{E}\xi_x^2 \in C_b(X)$, так как $G_n(\mathbf{E}\xi_x^2) \downarrow g(x)$ для любого $\mathbf{E}\xi_x^2 \in C_b(X)$.

3. Оптимизация распределения длины пробега

3.1. В качестве примера рассмотрим задачу об оценке вероятности вылета частиц из полупространства $-\infty < z \leq H$. Будем считать, что вне этого полупространства находится абсолютный поглотитель, причем средний свободный пробег σ^{-1} равен 1 во всем пространстве. Рассеяние частицы в точке столкновения описывается симметричной нормированной плотностью $w(\mu, \mu')$, где μ — косинус угла между направлением пробега и осью z. Вероятность выживания при столкновении в точке $z < H$ равна $q < 1$. Здесь $x = (z, \mu)$;

$$h(x) = \begin{cases} \exp(-(H-z)/\mu) & \text{при } z < H,\ \mu > 0, \\ h \equiv 0 & \text{иначе.} \end{cases}$$

Величина $\mathbf{E}\xi_x = \varphi^*(x)$ равна искомой вероятности вылета частицы, стартовавшей в точке x (см. [3, разд. 2.4]).

Рассмотрим оптимизацию распределения длины свободного пробега, физическая плотность распределения которой равна e^{-l}. При этом в уравнении (2.4) имеем $\mathbf{t} = (l, \mu)$, $z' = z + \mu l$ и

$$k_1(x, t_1') = e^{-l}, \quad k_2((x, t_1'), t_2') = q w(\mu, \mu') I\{z' < H\},$$

где $I\{S\}$ — индикатор множества S. Запишем следующее асимптотическое приближение (при $H \to \infty$) уравнения (2.4):

$$g(z, \mu) = \left(\int\limits_0^\infty e^{-l} \left(\int\limits_{-1}^1 q w(\mu, \mu') g(z'(z, \mu, l), \mu') d\mu' \right)^{1/2} dl \right)^2.$$

Будем искать решение этого уравнения в виде $g(z, \mu) = a(\mu) e^{2cz}$. Простые выкладки приводят к уравнению

$$a(\mu) = \frac{q}{(1 - c\mu)^2} \int\limits_{-1}^1 w(\mu, \mu') a(\mu')\, d\mu'. \tag{3.1}$$

Для его решения используем транспортное приближение [4]

$$w(\mu, \mu') = \mu_0 \delta(\mu' - \mu) + (1 - \mu_0)/2,$$

где μ_0 — средний косинус угла рассеяния. В силу однородности (3.1) можно предположить, что $\int\limits_{-1}^1 a(\mu)\, d\mu = 1$. Подставив в (3.1) транспортное приближение для $w(\mu, \mu')$, получим

$$a(\mu) = \frac{\frac{q}{2}(1 - \mu_0)}{(1 - c\mu)^2 - q\mu_0}.$$

Учитывая, что $\int\limits_{-1}^1 a(\mu)\, d\mu = 1$, приходим к следующему уравнению для критического значения c^*:

$$\ln \frac{1 - (c - (q\mu_0)^{1/2})^2}{1 - (c + (q\mu_0)^{1/2})^2} = c \frac{4}{1 - \mu_0} \left(\frac{\mu_0}{q} \right)^2.$$

В случае изотропного рассеяния из (3.1) имеем $c = (1-q)^{1/2}$.

3.2. Известно, что использование функции $g(z,\mu) = a(\mu)e^{cz}$ в (2.5) приводит к модификации распределения длины пробега, состоящей в замене $\sigma' = \sigma - c\mu$, т. е. к так называемому экспоненциональному преобразованию (см. [1]). При этом

$$\frac{k((x,\mathbf{t}),(x',\mathbf{t}'))}{p\ ((x,\mathbf{t}),(x',\mathbf{t}'))} = \frac{q\sigma e^{-lc\mu}}{(1-p)(\sigma - c\mu)},$$

где q — физическая вероятность выживания частицы при столкновении, а p — моделируемая вероятность обрыва траектории; в п. 3.1 рассматривалось $p = 1-q$. Функция $a(\mu)$ здесь роли не играет, так как сокращается в (2.5).

Для оптимизации экспоненциального преобразования можно пытаться использовать характеристическое уравнение переноса, определяющее параметр $-1/L$ пространственной экспонециальной асимптотики плотности потока частиц [5]; величина L называется диффузионной длиной. Соответствующая асимптотика (при $H \to \infty$) функции ценности имеет вид $C\exp(z/L)$. Однако численные эксперименты с экспоненциальным преобразованием показали, что при значениях параметра c, близких к решению характеристического уравнения переноса, соответствующая дисперсия существенно превосходит дисперсию оценок прямого моделирования. Из сказанного можно сделать вывод о неэффективности частичного «ценностного» моделирования, в котором функция ценности используется для модификации распределения только одной из вспомогательных случайных величин, в данном случае длины свободного пробега.

С другой стороны, ясно, что использование экспоненциального преобразования с параметром c_0, определяемым уравнением (3.1), дает существенное уменьшение величины $\mathbf{E}\xi_x^2$ сравнительно с прямым моделированием, причем асимптотику этого уменьшения по величине H можно оценить непосредственно на основе леммы 2.1. Величина $\mathbf{E}\xi_x^2$ при $z = H$, т. е. при нулевом расстоянии от источника до приемника, во всех вариантах моделирования несущественно отличается от единицы. Для оптимальной «частичной» модификации имеем

$$\mathbf{E}\xi_x^2 = g(x) = O(e^{-2c_0(H-z)}),$$

а для прямого моделирования

$$\mathbf{E}\xi_x^2 = \varphi^*(z,\mu)[1-\varphi^*(z,\mu)] = O(e^{-\frac{H-z}{L}}).$$

Следовательно, асимптотика уменьшения дисперсии определяется величиной $O(\exp\{-(H-z)(2c_0 - 1/L)\})$. Например, в случае изотропного рассеяния при $q = 0,7$ имеем $1/L = 0,829$, а $2c_0 = 2\sqrt{0.3} \approx 1.095$, т. е. $2c_0 - 1/L \approx 0.266$.

3.3. Поскольку $I_h = \mathbf{E}\big(\frac{f(x_0)}{\pi(x_0)}, \xi_{x_0}\big)$ и

$$\mathbf{E}\xi^2 = \mathbf{E}\xi^2(\pi) = \int\limits_X \frac{f^2(x)}{\pi(x)} \mathbf{E}\xi_x^2\, dx,$$

то минимум величины $\mathbf{D}\xi$, определяющий среднеквадратичную погрешность оценок линейного функционала J_h, достигается при

$$\pi(x) = \frac{f(x)\big(\mathbf{E}\xi_x^2\big)^{1/2}}{\int\limits_X f(x)\big(\mathbf{E}\xi_x^2\big)^{1/2}\, dx}.$$

Практически интересен вопрос об эффективности более простой, с точки зрения возможности приближенной оценки, начальной плотности

$$\pi^*(x) = f(x)\varphi^*(x)/I_h.$$

Следующий пример показывает, что такая плотность может быть малоэффективна.

Предположим, что величина ξ_x является «бернуллиевой», т. е.

$$P(\xi_x = 1) = p(x), \quad P(\xi_x = 0) = 1 - p(x), \quad 0 < p(x) < 1.$$

В этом случае

$$\mathbf{E}\xi_x^2 = p(x) = \varphi^*(x), \quad \mathbf{E}\xi^2(\pi^*) = \mathbf{E}\xi^2(f) = I_h.$$

Таким образом, «ценностное» моделирование начальной точки траектории для «бернуллиевого» варианта ξ_x, например, когда просто подсчитывается число поглощенных частиц, не уменьшает дисперсию сравнительно с прямым моделированием.

ЛИТЕРАТУРА

1. *Ермаков С. М., Михайлов Г. А.* Статистическое моделирование. М.: Наука, 1982.

2. *Фролов А. С., Ченцов Н. Н.* Решение трех типичных задач теории переноса методом Монте-Карло // Метод Монте-Карло в проблеме переноса излучений. М.: Атомиздат, 1967.

3. *Михайлов Г. А.* Оптимизация весовых методов Монте-Карло. М.: Наука, 1987.

4. *Марчук Г. И., Михайлов Г. А., Назаралиев М. А. и др.* Метод Монте-Карло в атмосферной оптике. Новосибирск: Наука, 1976.

5. *Дэвисон Б.* Теория переноса нейтронов. М.: Атомиздат, 1960.

Статья поступила 25 марта 2003 г.

Михайлов Геннадий Алексеевич, Медведев Илья Николаевич
Институт вычислительной математики и математической геофизики СО РАН, пр. Академика М. А. Лаврентьева, 6, Новосибирск 630090
gam@sscc.ru, medvedev@ngs.ru