

Векторные алгоритмы метода Монте-Карло с конечной трудоемкостью ¹

И.Н. Медведев

Институт вычислительной математики и математической геофизики СО РАН (Новосибирск),

Новосибирский государственный университет (Новосибирск)

e-mail: min@osmf.sscs.ru

Аннотация.

В работе изучается вопрос конечности трудоемкости весовых алгоритмов метода Монте-Карло для оценки линейных функционалов от решения системы интегральных уравнений 2-го рода. Построена универсальная модификация весовой векторной оценки по столкновениям с ветвлением траектории цепи соответственно элементам матричного веса. Доказано, что трудоемкость построенного алгоритма ограничена, если ограничены базовые функционалы. Представлены результаты численных расчетов с использованием модифицированной весовой оценки для одной задачи теории переноса излучения с учетом поляризации.

Ключевые слова: система линейных интегральных уравнений 2-го рода, весовая векторная оценка, матричный вес, ветвление траектории в цепи Маркова, конечная трудоемкость алгоритма, перенос поляризованного излучения.

1. Вводная информация

Рассмотрим систему интегральных уравнений второго рода:

$$\varphi_i(x) = \sum_{j=1}^m \int_X k_{ij}(x, y) \varphi_j(y) dy + h_i(x) \quad (1.1)$$

или в векторном виде $\Phi = \mathbf{K}\Phi + H$, где $H^T = (h_1, \dots, h_m)$,

$$\mathbf{K} \in [L_\infty \rightarrow L_\infty], \quad \|H\|_{L_\infty} = \text{vrai} \sup_{i,x} |h_i(x)|,$$

а интегрирование производится по мере Лебега в κ -мерном евклидовом пространстве X .

Предполагается, что спектральный радиус $\lambda(\mathbf{K}) < 1$. При этом имеет место разложение решения в ряд Неймана

$$\Phi = \sum_{n=0}^{\infty} \mathbf{K}^n H. \quad (1.2)$$

Поскольку здесь

$$\lambda(\mathbf{K}) = \underline{\lim} \|\mathbf{K}^n\|_{L_\infty}^{1/n} = \inf \|\mathbf{K}^n\|_{L_\infty}^{1/n},$$

то для сходимости ряда (1.2) достаточно выполнения неравенства $\|\mathbf{K}^{n_0}\| < 1$ для некоторого $n_0 \geq 1$. Отметим, что здесь

¹ Работа выполнена при финансовой поддержке РФФИ (проекты 15-01-00894 а, 16-01-00530 а) и программы фундаментальных исследований Президиума РАН I.33.

$$\|\mathbf{K}\| = \sup_{x,i} \sum_{j=1}^m \int |k_{ij}(x,y)| dy.$$

Рассмотрим цепь Маркова $\{x_n\}$, ($n = 0, \dots, N$) с плотностью перехода $p(x, y)$, причем величина

$$p(x) = 1 - \int_X p(x, y) dy \geq 0$$

рассматривается как вероятность обрыва (остановки) в точке x , N -случайный номер последнего состояния $x_0 \equiv x$. Известно [2], что такая цепь обрывается с вероятностью 1 и более того, $E(N) < +\infty$, если $\lambda(B_p) < 1$, где B_p – интегральный оператор с ядром $p(x, y)$ (в частности, если $p(x) \geq \varepsilon > 0$).

Стандартная векторная оценка метода Монте-Карло “по столкновениям” для величины $\Phi(x)$ строится на основе соотношений

$$\Phi(x) = E\xi_x, \quad \xi_x = H(x) + \sum_{n=1}^N Q_n H(x_n),$$

$$Q_0 = I, \quad Q_{n+1} = Q_n K(x_n, x_{n+1}) / p(x_n, x_{n+1}), \quad n = 0, 1, \dots,$$

где I – единичная матрица, а $K(x, y)$ – матрица ядер $\{k_{ij}(x, y)\}$, ($i, j = 1, \dots, m$). Заметим, что соотношение $\Phi(x) = E\xi_x$ справедливо при выполнении аналогичных рассмотренным в [2] для случая $m = 1$ “условий несмещенности” и дополнительного условия $\lambda(\mathbf{K}_1) < 1$, где \mathbf{K}_1 – оператор, получаемый из оператора \mathbf{K} заменой ядер на их модули. Здесь условия несмещенности [2] преобразуются к виду:

$$p(x, y) > 0, \quad \text{если} \quad \sum_{i,j=1}^m |k_{ij}(x, y)| > 0, \quad (1.3)$$

а оценка ξ_x однозначно определяется рекурсией

$$\xi_x = H(x) + \delta_y Q(x, y) \xi_y, \quad (1.4)$$

где $Q(x, y) = K(x, y) / p(x, y)$ – матричный вес, а δ_y – индикатор необрыва цепи при переходе $x \rightarrow y$.

В [2] приведено следующее уравнение для матрицы вторых моментов $\Psi(x) = E(\xi_x \xi_x^T)$:

$$\Psi(x) = \chi(x) + \int_X \frac{K(x, y) \Psi(y) K^T(x, y)}{p(x, y)} dy, \quad (1.5)$$

или $\Psi = \chi + \mathbf{K}_p \Psi$, где

$$\chi = H \Phi^T + \Phi H^T - H H^T.$$

Это уравнение рассматривается в пространстве \mathbf{L}_∞ матричнозначных функций с нормой

$$\|\Psi\| = \text{vrai sup}_{i,j,x} |\Psi_{i,j}(x)|.$$

и имеет место [2] следующее выражение:

$$\|\mathbf{K}_p\|_{\mathbf{L}_\infty} = \sup_{i,x} \int \frac{\left(\sum_{j=1}^m |k_{ij}(x,y)| \right)^2}{p(x,y)} dy. \quad (1.6)$$

Оператор, получаемый из \mathbf{K}_p заменой ядер на их модули, обозначим $\mathbf{K}_{p,1}$. Предполагается, что $\mathbf{K}_{p,1} \in [\mathbf{L}_\infty \rightarrow \mathbf{L}_\infty]$ и, следовательно, \mathbf{K}_p обладает тем же свойством.

В работе [6], с использованием разработанного авторами метода рекуррентного “частичного” осреднения [3], было доказано следующее утверждение.

Теорема 1. *Если $\lambda(\mathbf{K}_1) < 1$ и $\lambda(\mathbf{K}_{p,1}) < 1$, то $\Psi(x) = \mathbb{E}(\boldsymbol{\xi}_x \boldsymbol{\xi}_x^\top)$ является решением уравнения (1.5) и $\Psi \in \mathbf{L}_\infty$.*

Методом Монте-Карло обычно оценивают линейные функционалы вида [1]

$$I = (F, \Phi) = \int_X F^\top(x) \Phi(x) dx, \quad (1.7)$$

где $F^\top(x) = (f_1(x), \dots, f_m(x))$, причем

$$\|F^\top\|_{L_1} = \sum_{j=1}^m \int_X |f_j(x)| dx < \infty.$$

Пусть точка x_0 распределена с плотностью вероятностей $\pi(x)$ такой, что $\pi(x) \neq 0$, если $F^\top(x) \Phi(x) \neq 0$. Тогда, полагая $\boldsymbol{\xi} = F^\top(x_0) \boldsymbol{\xi}_{x_0} / \pi(x_0)$, имеем

$$I = \mathbb{E} \boldsymbol{\xi} = \mathbb{E} \left\{ \frac{F^\top(x_0)}{\pi(x_0)} \boldsymbol{\xi}_{x_0} \right\},$$

$$\mathbb{E} \boldsymbol{\xi}^2 = \mathbb{E} \left\{ \frac{F^\top(x_0) \boldsymbol{\xi}_{x_0} \boldsymbol{\xi}_{x_0}^\top F(x_0)}{\pi^2(x_0)} \right\} = \mathbb{E} \left\{ \frac{F^\top(x_0) \Psi(x_0) F(x_0)}{\pi^2(x_0)} \right\}. \quad (1.8)$$

Таким образом, дисперсия $D\boldsymbol{\xi}$ определяется матрицей вторых моментов $\Psi(x)$. В частности, $D\boldsymbol{\xi} < +\infty$, если $\lambda(\mathbf{K}_{p,1}) < 1$ и $F^\top(x)/\pi(x) \in L_1(X)$ [6]. Если спектральный радиус $\lambda(\mathbf{K}_{p,1}) > 1$, то величина $D\boldsymbol{\xi}$ может быть бесконечно большой и использование весовой оценки $\boldsymbol{\xi}$ для вычисления функционала I нецелесообразно.

2. Векторная оценка с ветвлением траектории

Рассмотрим систему уравнений (1.1) с неотрицательными компонентами $k_{ij}(x,y), h_i(x), (i, j = 1, \dots, m)$. Введем целочисленную неотрицательную случайную величину (число “ветвей”) ν_n с заданным распределением вероятностей. Определим случайную векторную оценку с ветвлением траектории следующей рекурсией:

$$\boldsymbol{\zeta}_{x_0} = H(x_0) + \delta_{x_1} \frac{Q(x_0, x_1)}{\mathbb{E} \nu_1} \sum_{i=1}^{\nu_1} \boldsymbol{\zeta}_{x_1}^{(i)}, \quad (2.1)$$

$$\boldsymbol{\zeta}_{x_{n-1}} = H(x_{n-1}) + \delta_{x_n} \frac{Q(x_{n-1}, x_n)}{\mathbb{E} \nu_n} \sum_{i=1}^{\nu_n} \boldsymbol{\zeta}_{x_n}^{(i)}, \quad (2.2)$$

где $\boldsymbol{\zeta}_{x_n}^{(\cdot)}$ – независимые реализации $\boldsymbol{\zeta}_{x_{n-1}}$.

Лемма 1. При выполнении условий несмещенности (1.3) и $\lambda(\mathbf{K}) < 1$, имеет место соотношение $E\zeta_x = \Phi(x)$.

Доказательство. Так как все элементы в выражениях (2.1), (2.2) неотрицательны, то $\forall n \geq 1$ в силу тождества Вальда имеем

$$E \sum_{i=1}^{\nu_n} \zeta_{x_n}^{(i)} = E\nu_n E\zeta_{x_n}.$$

Это равенство выполняется и в случае $E\zeta_{x_n} = +\infty$, так как вследствие неотрицательности элементов задачи имеем:

$$E \sum_{i=1}^{\nu_n} \zeta_{x_n}^{(i)} = EE \left(\sum_{n=1}^{\nu} \zeta_{x_n}^{(i)} | \nu \right).$$

Следовательно, для вычисления величины $E\zeta_{x_0}$ можно последовательно применять рекурсию вида (1.4). \square

Согласно рекурсии (2.2), после успешного перехода $x_{n-1} \rightarrow x_n$ из фиксированной вершины “дерева” в n -ом поколении “испускаются” ν_n независимых траекторий с “полным” матричным весом

$$\tilde{Q}_n = \frac{Q(x_0, x_1)Q(x_1, x_2)\dots Q(x_{n-1}, x_n)}{E\nu_1 \dots E\nu_n} = \frac{Q_{n-1}Q(x_{n-1}, x_n)}{E\nu_1 \dots E\nu_{n-1}E\nu_n} \quad (n \geq 1). \quad (2.3)$$

Введем обозначение для произвольного элемента $\{A\}_{ij} = a_{ij}$ матрицы A , определим последовательность среднего числа ветвей $E\nu_n$

$$E\nu_1 = \sup_i \sum_{j=1}^m \{Q(x_0, x_1)\}_{ij}, \quad E\nu_n = \sup_i \sum_{j=1}^m \left\{ \frac{Q_{n-1}Q(x_{n-1}, x_n)}{E\nu_1 \dots E\nu_{n-1}} \right\}_{ij}, \quad (n \geq 2), \quad (2.4)$$

и прямой подстановкой проверим, что

$$E\nu_1 E\nu_2 \dots E\nu_n = \sup_i \sum_{j=1}^m \{Q_{n-1}Q(x_{n-1}, x_n)\}_{ij}. \quad (2.5)$$

Здесь и далее будем полагать, что величина ν_n имеет следующее распределение вероятностей

$$P(\nu_n = [E\nu_n]) = 1 - (E\nu_n - [E\nu_n]), \quad P(\nu_n = [E\nu_n] + 1) = E\nu_n - [E\nu_n]. \quad (2.6)$$

Нетрудно проверить, что распределение (2.6) определяет минимальное значение $D\nu$ в классе случайных целочисленных величин с фиксированным значением $E\nu_n$ [3] и

$$E\nu_n^2 = [E\nu_n]^2 + (2[E\nu_n] + 1)(E\nu_n - [E\nu_n]). \quad (2.7)$$

Теорема 2. Если выполняются условия леммы 1, то $\tilde{\Psi}(x) = E(\zeta_x \zeta_x^T) \in \mathbf{L}_\infty(X)$.

Доказательство. Как и в работе [7], можно проверить, что для любого $n \geq 1$

$$\zeta_{x_{n-1}} \zeta_{x_{n-1}}^T =$$

$$\begin{aligned}
& H(x_{n-1})H^T(x_{n-1}) + \delta_{x_n} \frac{Q(x_{n-1}, x_n)}{E\nu_n} \sum_{i=1}^{\nu_n} \zeta_{x_n}^{(i)} H^T(x_n) + \delta_{x_n} H(x_n) \left(\sum_{i=1}^{\nu_n} \zeta_{x_n}^{(i)} \right)^T \frac{Q^T(x_{n-1}, x_n)}{E\nu_n} \\
& + \delta_{x_n} \frac{Q(x_{n-1}, x_n)}{E\nu_n} \sum_{i=1}^{\nu_n} \zeta_{x_n}^{(i)} \left(\sum_{i=1}^{\nu_n} \zeta_{x_n}^{(i)} \right)^T \frac{Q^T(x_{n-1}, x_n)}{E\nu_n}.
\end{aligned}$$

Так как по определению компоненты $\zeta_{x_n}^{(i)}$ независимы, следовательно, в силу тождества Вальда имеем

$$E \left(\sum_{i=1}^{\nu_n} \zeta_{x_n}^{(i)} \right) = E\nu_n E\zeta_{x_n},$$

$$E \left(\sum_{i=1}^{\nu_n} \zeta_{x_n}^{(i)} \right) \left(\sum_{i=1}^{\nu_n} \zeta_{x_n}^{(i)} \right)^T = E\nu_n E\zeta_{x_n} \zeta_{x_n}^T + E(\nu_n(\nu_n - 1)) \Phi(x_n) \Phi^T(x_n).$$

С учетом последних равенств для математических ожиданий получаем, что величина $\tilde{\Psi}(x_n) = E(\zeta_{x_n} \zeta_{x_n}^T)$ для любого $n \geq 1$ удовлетворяет интегральному уравнению:

$$\tilde{\Psi}(x_{n-1}) = \tilde{\chi}_{\nu_n}(x_{n-1}) + \int_X \frac{K(x_{n-1}, x_n) \Psi(x_n) K^T(x_{n-1}, x_n)}{E\nu_n p(x_{n-1}, x_n)} dx_n,$$

или

$$\tilde{\Psi}(x_{n-1}) = \tilde{\chi}_{\nu_n}(x_{n-1}) + [\mathbf{K}_{p, \nu_n} \tilde{\Psi}](x_{n-1}). \quad (2.8)$$

где

$$\tilde{\chi}_{\nu_n}(x_{n-1}) = \chi(x_{n-1}) + \int_X \frac{K(x_{n-1}, x_n) E(\nu_n(\nu_n - 1)) \Phi(x_n) \Phi^T(x_n) K^T(x_{n-1}, x_n)}{(E\nu_n)^2 p(x_{n-1}, x_n)} dx_n.$$

Заметим, что второе слагаемое в выражении для $\tilde{\chi}_{\nu_n}$ имеет место только когда $E\nu_n > 1$. Учитывая последнее замечание и выражение (2.7), легко проверить, что $\forall n \geq 1$

$$\frac{E(\nu_n(\nu_n - 1))}{(E\nu_n)^2} \leq \frac{[E\nu_n]^2 + 2([E\nu_n] + 1)}{[E\nu_n]^2} \leq \frac{([E\nu_n] + 1)^2 + 1}{[E\nu_n]^2} \leq 3,$$

и, согласно (1.5),

$$\|\tilde{\chi}_{\nu_n}\| \leq \|H\Phi^T + \Phi H^T - HH^T\| + 3\|\mathbf{K}_p \Phi \Phi^T\| < \infty \quad (2.9)$$

в силу того, что $\lambda(\mathbf{K}_1) < 1$ и $\mathbf{K}_p \in [\mathbf{L}_\infty \rightarrow \mathbf{L}_\infty]$.

Пусть $\psi_{n-1}(x_{n-1}) = \tilde{\Psi}(x_{n-1})$. Тогда, для любого $n \geq 1$, соотношение (2.8) можно представить в операторном виде

$$\psi_{n-1} = \tilde{\chi}_{\nu_n} + \mathbf{K}_{p, \nu_n} \psi_n \quad (n \geq 1),$$

и прямой подстановкой можно проверить, что

$$\begin{aligned}
\psi_{n-1} &= \tilde{\chi}_{\nu_n} + \sum_{i=0}^{\infty} \mathbf{K}_{p, \nu_n} \mathbf{K}_{p, \nu_{n+1}} \cdots \mathbf{K}_{p, \nu_{n+i}} \tilde{\chi}_{\nu_{n+i+1}} \\
&= \tilde{\chi}_{\nu_n} + \mathbf{K}_{p, \nu_n} (\tilde{\chi}_{\nu_{n+1}} + \sum_{i=0}^{\infty} \mathbf{K}_{p, \nu_{n+1}} \mathbf{K}_{p, \nu_{n+2}} \cdots \mathbf{K}_{p, \nu_{n+1+i}} \tilde{\chi}_{\nu_{n+1+i+1}}) = \tilde{\chi}_{\nu_n} + \mathbf{K}_{p, \nu_n} \psi_n.
\end{aligned}$$

Следовательно,

$$\begin{aligned}
\tilde{\Psi}(x_0) &= \psi_0(x_0) = \tilde{\chi}_{\nu_1}(x_0) + \sum_{i=0}^{\infty} [\mathbf{K}_{p,\nu_1} \mathbf{K}_{p,\nu_2} \dots \mathbf{K}_{p,\nu_{i+1}} \tilde{\chi}_{\nu_{i+1}}](x_0) \\
&= \tilde{\chi}_{\nu_1}(x_0) + \sum_{n=1}^{\infty} [\mathbf{K}_{p,\nu_1} \mathbf{K}_{p,\nu_2} \dots \mathbf{K}_{p,\nu_n} \tilde{\chi}_{\nu_{n+1}}](x_0),
\end{aligned} \tag{2.10}$$

где оператор $\mathbf{K}_{p,\nu_1} \mathbf{K}_{p,\nu_2} \dots \mathbf{K}_{p,\nu_n}$ для произвольной $D \in \mathbf{L}_{\infty}$ и $n \geq 1$ определяется выражением:

$$\begin{aligned}
&[\mathbf{K}_{p,\nu_1} \mathbf{K}_{p,\nu_2} \dots \mathbf{K}_{p,\nu_n} D](x_0) \\
&= \int_X \dots \int_X \frac{K(x_0, x_1) \dots K(x_{n-1}, x_n) D(x_n) K^T(x_{n-1}, x_n) \dots K^T(x_0, x_1)}{\mathbf{E}\nu_1 \mathbf{E}\nu_2 \dots \mathbf{E}\nu_n p(x_0, x_1) \dots p(x_{n-1}, x_n)} dx_1 \dots dx_n.
\end{aligned}$$

Согласно выражениям (1.6), (2.5), справедливо следующее неравенство:

$$\begin{aligned}
&\|\mathbf{K}_{p,\nu_1} \mathbf{K}_{p,\nu_2} \dots \mathbf{K}_{p,\nu_n}\| \\
&= \sup_{i, x_0} \int_X \dots \int_X \frac{\left(\sum_{j=1}^m \{K(x_0, x_1) \dots K(x_{n-1}, x_n)\}_{ij} \right)^2}{\mathbf{E}\nu_1 \dots \mathbf{E}\nu_n p(x_0, x_1) \dots p(x_{n-1}, x_n)} dx_1 \dots dx_n \\
&\leq \sup_{i, x_0} \int_X \dots \int_X \sum_{j=1}^m \{K(x_0, x_1) \dots K(x_{n-1}, x_n)\}_{ij} dx_1 \dots dx_n = \|\mathbf{K}_1^n\|.
\end{aligned}$$

Из последнего неравенства и условий $\lambda(\mathbf{K}) < 1$, (2.9), следует, что ряд Неймана (2.10) для матрицы вторых моментов $\tilde{\Psi}(x) = \mathbf{E}(\zeta_x \zeta_x^T)$ сходится. \square

Известно, что трудоемкость методов Монте-Карло определяется величиной $S_{\eta} = T_{\eta} D \eta$, где T_{η} - среднее время моделирования на ЭВМ для получения одного выборочного значения η [5], [2]. В свою очередь, величина T_{η} пропорциональна среднему числу $\mathbf{E}N_{\eta}$ столкновений в цепи Маркова для получения одного выборочного значения η . Для векторной оценки (2.2) с ветвлением траектории согласно (2.4), число N_{ζ} определяется следующей рекурсией:

$$N_{\zeta} = N_{\zeta_{x_0}}, \quad N_{\zeta_{x_0}} = 1 + \delta_{x_1} \sum_{i=1}^{\nu_1} N_{\zeta_{x_1}}^{(i)}, \quad N_{\zeta_{x_{n-1}}} = 1 + \delta_{x_n} \sum_{i=1}^{\nu_n} N_{\zeta_{x_n}}^{(i)}, \tag{2.11}$$

где $N_{\zeta_{x_n}}^{(i)}$ - независимые реализации $N_{\zeta_{x_n}}$.

Теорема 3. Величина $\mathbf{E}N_{\zeta}$ ограничена, если $\lambda(\mathbf{K}) < 1$.

Доказательство. Методом “частичного” осреднения [3] рекурсии (2.11) можно проверить, что неотрицательная величина $\mathbf{E}N_{\zeta_{x_n}} = g(x_n)$ для любого $n \geq 1$ удовлетворяет интегральному соотношению:

$$g(x_{n-1}) = 1 + \int_X p(x_{n-1}, x_n) \mathbf{E}\nu_n g(x_n) dx_n, \quad \text{или} \quad g(x_{n-1}) = 1 + [B_{p_n, \nu_n} g](x_{n-1}).$$

Пусть $G_{n-1}(x_{n-1}) = g(x_{n-1})$. Тогда, для любого $n \geq 1$, последнее интегральное соотношение (2.8) можно представить в операторном виде

$$G_{n-1} = 1 + B_{p,\nu_n} G_n \quad (n \geq 1),$$

и прямой подстановкой можно проверить, что

$$\begin{aligned} G_{n-1} &= 1 + \sum_{i=0}^{\infty} B_{p,\nu_n} B_{p,\nu_{n+1}} \dots B_{p,\nu_{n+i}} 1 \\ &= 1 + B_{p,\nu_n} (1 + \sum_{i=0}^{\infty} B_{p,\nu_{n+1}} B_{p,\nu_{n+2}} \dots B_{p,\nu_{n+1+i}} 1) = 1 + B_{p,\nu_n} G_n. \end{aligned}$$

Следовательно,

$$\begin{aligned} g(x_0) = G_0(x_0) &= 1 + \sum_{i=0}^{\infty} [B_{p,\nu_1} B_{p,\nu_2} \dots B_{p,\nu_{1+i}}](x_0) \quad (2.12) \\ &= 1 + \sum_{n=1}^{\infty} [B_{p,\nu_1} B_{p,\nu_2} \dots B_{p,\nu_n} 1](x_0), \end{aligned}$$

где оператор $B_{p,\nu_1} B_{p,\nu_2} \dots B_{p,\nu_n}$ для произвольной $u \in L_{\infty}$ при $n \geq 1$ с учетом (2.4), (2.5), определяется следующим выражением

$$\begin{aligned} &[B_{p,\nu_1} B_{p,\nu_2} \dots B_{p,\nu_n} u](x_0) \\ &= \int_X \dots \int_X p(x_0, x_1) E\nu_1 \dots p(x_{n-1}, x_n) E\nu_n u(x_n) dx_1 \dots dx_n \\ &= \int_X \dots \int_X p(x_0, x_1) \dots p(x_{n-1}, x_n) E\nu_1 \dots E\nu_n u(x_n) dx_1 \dots dx_n \\ &= \int_X \dots \int_X \sup_{x_0, i} \sum_{j=1}^m |\{K(x_0, x_1) \dots K(x_{n-1}, x_n)\}_{ij}| u(x_n) dx_1 \dots dx_n, \end{aligned}$$

причем

$$\begin{aligned} \|B_{p,\nu_1} B_{p,\nu_2} \dots B_{p,\nu_n}\| &= \int_X \dots \int_X \sup_{x_0, i} \sum_{j=1}^m |\{K(x_0, x_1) \dots K(x_{n-1}, x_n)\}_{ij}| dx_1 \dots dx_n \\ &\leq \int_X \dots \int_X \sum_{i=1}^m \sum_{j=1}^m |\{K(x_0, x_1) \dots K(x_{n-1}, x_n)\}_{ij}| dx_1 \dots dx_n \\ &\leq \sum_{i=1}^m \int_X \dots \int_X \sum_{j=1}^m |\{K(x_0, x_1) \dots K(x_{n-1}, x_n)\}_{ij}| dx_1 \dots dx_n \\ &\leq m \sup_{x_0, i} \int_X \dots \int_X \sum_{j=1}^m |\{K(x_0, x_1) \dots K(x_{n-1}, x_n)\}_{ij}| dx_1 \dots dx_n = m \|\mathbf{K}_1^n\| \end{aligned}$$

Из последнего неравенства и условия $\lambda(\mathbf{K}) < 1$, следует, что ряд Неймана (2.12) для величины $E N_{\zeta}$ сходится. \square

В конце раздела 1 было отмечено, что существенным ограничением на использование стандартной векторной оценки $\boldsymbol{\xi} = F^T(x_0)\boldsymbol{\xi}_{x_0}/\pi(x_0)$ для вычисления функционала $I = (F, \Phi)$ является условие $\lambda(\mathbf{K}_{p,1}) < 1$. Для построенной векторной оценки с ветвлением траектории $\boldsymbol{\zeta} = F^T(x_0)\boldsymbol{\zeta}_{x_0}/\pi(x_0)$, в силу теорем 2, 3, справедливо более слабое ограничение.

Утверждение 1 *Трудоёмкость S_{ζ} ограничена, если ограничен исходный функционал I ($\lambda(\mathbf{K}) < 1$ и $F^T(x)/\pi(x) \in L_1$).*

3. Приложение в теории переноса излучения с учетом поляризации

Известно, что для описания поляризационных свойств света удобно использовать вектор Стокса [1, 2]

$$\tilde{\Phi}(x) = (\tilde{\varphi}_1(x), \tilde{\varphi}_2(x), \tilde{\varphi}_3(x), \tilde{\varphi}_4(x))^T,$$

причем, вектор функции $\tilde{\Phi}$ образуют конус S_t , определяемый соотношениями:

$$\tilde{\varphi}_1(x) \geq 0, \quad \tilde{\varphi}_2^2(x) + \tilde{\varphi}_3^2(x) + \tilde{\varphi}_4^2(x) \leq \tilde{\varphi}_1^2(x), \quad (3.1)$$

а компоненты вектора $\tilde{\Phi}$ удовлетворяют системе интегральных уравнений переноса с учетом поляризации вида (1.1):

$$\tilde{\varphi}_i(x) = \int \sum_{j=1}^4 k_{ij}(x, x') \tilde{\varphi}_j(x') dx' + h_i(x), \quad i = 1, \dots, 4, \quad \text{или} \quad \tilde{\Phi} = \mathbf{K}\tilde{\Phi} + H \quad (3.2)$$

Здесь $x = (\mathbf{r}, \omega)$, где \mathbf{r} – точка физического пространства R , $\omega \in \Omega$ – единичный вектор направления пробега частицы и $\mathbf{K} \in [S_t \rightarrow S_t]$ [2]. Введем следующие обозначения: $\mu = (\omega, \omega')$ – косинус угла рассеяния, θ – азимутальный угол рассеяния, $p_2(\mu)$ – индикатриса рассеяния, $\sigma(\mathbf{r}) = \sigma_s(\mathbf{r}) + \sigma_c(\mathbf{r})$ – полное сечение, $\sigma_s(\mathbf{r})$ и $\sigma_c(\mathbf{r})$ – сечения рассеяния и поглощения соответственно, $q(\mathbf{r}) = \sigma_c(\mathbf{r})/\sigma(\mathbf{r})$ – вероятность рассеяния, l – длина свободного пробега, $p_\chi(l; \mathbf{r}, \omega')$ – субстохастическая плотность распределения длины пробега из точки \mathbf{r} в направлении ω' . Известно, что

$$p_\chi(l; \mathbf{r}, \omega') = \sigma(\mathbf{r} + \omega'l) \exp\left(-\int_0^l \sigma(\mathbf{r} + s\omega') ds\right), \quad l \leq l^*(l\mathbf{r}, \omega'),$$

где $l^*(\mathbf{r}, \omega')$ – расстояние от точки \mathbf{r} вдоль направления ω' до границы среды, которую можно считать выпуклой. С учетом введенных вспомогательных переменных μ, θ, l , матрица ядер для системы (3.2) задается соотношением [1],[4]:

$$K(x, y) = qp_\chi(l; \mathbf{r}, \omega') P^T(\mu, \theta) \delta(\omega' - \omega'(\omega, \mu, \theta)) \delta(\mathbf{r}' - \mathbf{r} - \omega'l), \quad (3.3)$$

где $y = (\mathbf{t}', x') = (\mu, \theta, l, x')$, $P^T(\mu, \theta) = L(i_1)R^T(\mu)L(-\pi + i_2)/2\pi$,

$$R(\mu) = \begin{pmatrix} r_{11} & r_{12} & 0 & 0 \\ r_{21} & r_{22} & 0 & 0 \\ 0 & 0 & r_{33} & r_{34} \\ 0 & 0 & -r_{43} & r_{44} \end{pmatrix}, \quad L(i) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2i & \sin 2i & 0 \\ 0 & -\sin 2i & \cos 2i & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

где $i_k = i_k(\omega, \mu, \theta)$; $k = 1, 2$; $\theta \in U(0, 2\pi)$; $r_{ij} = r_{ij}(\mu)$; $r_{11} \geq 0$; $\int_{-1}^{+1} r_{11}(\mu) d\mu = 1$. Предполагается, что среда изотропна и P не зависит от \mathbf{r} . Если рассеивающие частицы сами являются однородными сферами, то $r_{11} = r_{22}, r_{12} = r_{21}, r_{33} = r_{44}, r_{34} = r_{43}$ [1],[2].

Известно [1],[2], что для оценки линейных функционалов от решения системы (3.2) методом Монте-Карло можно построить весовую векторную оценку по столкновениям ξ_x вида (1.4). Несмотря на знакопеременность ядер матрицы $K(x, y)$, в работах [1],[2] было доказано, что, в силу свойств вектор-функций Стокса (3.1) и оператора $\mathbf{K} \in [S_t \rightarrow S_t]$ из (3.2), для выполнения равенства $E\xi_x = \tilde{\Phi}(x)$ достаточно, чтобы выполнялось условие $\lambda(\mathbf{K}) < 1$ - более слабое, чем в разделе 1. Если используется переходная плотность вида

$$p(x, y) = q_1 p_\chi^{(1)}(l; \mathbf{r}, \omega') p_2(\mu) \delta(\omega' - \omega'(\omega, \mu, \varphi)) \delta(\mathbf{r}' - \mathbf{r} - \omega'l)/(2\pi), \quad (3.4)$$

то, по сравнению с условиями теоремы 1, здесь для ограниченности элементов матрицы вторых моментов $E(\xi_x \xi_x^T)$ достаточно потребовать, чтобы выполнялось условие: $\lambda(\mathbf{K}_p) < 1$ ([2], [4]). В работе [2] было получено следующее неравенство:

$$\lambda(\mathbf{K}_p) \leq q_0 \lambda(S_p), \quad \text{где} \quad q_0 = \sup_{\mathbf{r}, \omega} \int_0^\infty \frac{q^2 p_\chi^2(l; \mathbf{r}, \omega)}{q_1 p_\chi^{(1)}(l; \mathbf{r}, \omega)} dl, \quad (3.5)$$

а S_p - оператор, получаемый из \mathbf{K}_p подстановкой $x \rightarrow \omega$, $y \rightarrow \omega'$, $p \rightarrow p_2/2\pi$, $K \rightarrow P^T$ (S_p соответствует "чистому" рассеянию в бесконечно однородной среде).

Таким образом, если $q_0 < 1/\lambda(S_p)$, то обычно используемые оценки вида $\xi = F^T(x_0)\xi_{x_0}/\pi(x_0)$ имеют конечную дисперсию при надлежащем выборе $\pi(x)$ [2]. В частности, для молекулярного рассеяния [1]

$$r_{11}(\mu) = \frac{3(1 + \mu^2)}{8}, \quad r_{12}(\mu) = -\frac{3(1 - \mu^2)}{8}, \quad r_{33}(\mu) = \frac{3\mu}{4}, \quad r_{34}(\mu) = 0, \quad \mu \in [-1, 1], \quad (3.6)$$

при $p_2 \equiv r_{11}$ было вычислено значение $\lambda(S_p) = 1 + (3\pi - 8)/8 \approx 1.178$ [4], а следовательно, дисперсия оценки ξ в реальной среде конечна только при достаточно большом поглощении. Например, при "физическом" моделировании ($q_1 \equiv q, p_\chi^{(1)} = p_\chi$) в бесконечно однородной среде из (3.5) легко получить неравенство

$$\lambda(\mathbf{K}_p) \leq (1 - \sigma_c/\sigma)\lambda(S_p)$$

и $\lambda(\mathbf{K}_p) < 1$ при $\sigma_c/\sigma > 0.151$. Существенно уменьшить величину σ_c/σ можно модификацией процесса переноса путем замены $\sigma \rightarrow \sigma_s$, $\sigma_c \rightarrow 0$ ([1]). В этой модификации поглощение учитывается соответствующим весовым множителем [1] и из неравенства (3.5) следует, что

$$\lambda(\mathbf{K}_p) \leq \frac{1 - \sigma_c/\sigma}{1 + \sigma_c/\sigma} \lambda(S_p)$$

и $\lambda(\mathbf{K}_p) < 1$ при $\sigma_c/\sigma > 0.082$. Даже при моделировании "без поглощения" [1],[5] ($q_1 \equiv 1, p_\chi^{(1)} = p_\chi$) получаем, что

$$\lambda(\mathbf{K}_p) \leq (1 - \sigma_c/\sigma)^2 \lambda(S_p)$$

и $\lambda(\mathbf{K}_p) < 1$ при $\sigma_c/\sigma > 0.0787$.

Таким образом при $\sigma_c/\sigma < 0.0787$ дисперсия векторной оценки может быть бесконечно большой величиной и вопрос об обоснованности применения векторного алгоритма Монте-Карло для оценки решения системы (3.2) в случае молекулярного рассеяния

остаётся открытым. В этом случае для оценки линейных функционалов от решения системы (3.2) целесообразно использовать векторную весовую оценку ζ с ветвлением траектории (см. раздел 2), для которой среднее количество ветвей после каждого перехода

$$E\nu_n = \sup_{x_0, i} \sum_{j=1}^m \left| \left\{ \frac{Q_{n-1}}{E\tilde{\nu}_{n-1}} Q(x_{n-1}, x_n) \right\}_{ij} \right| \quad (3.7)$$

будет определяться уже знакопеременными элементами оператора \mathbf{K} . Используя свойства вектор-функций Стокса (3.1) и оператора $\mathbf{K} \in [S_t \rightarrow S_t]$, можно проверить, что утверждения леммы 1 и теорем 2, 3 для оценки ζ также имеют место при условии $\lambda(\mathbf{K}) < 1$, т.е. величина $D\zeta$ и трудоемкость векторного алгоритма с ветвлением траектории конечны.

4. Тестовая задача

Рассмотрим бесконечную однородную среду, заполненную рассеивающим и поглощающим свет веществом, с источником излучения в точке $x_0 = (\mathbf{r}_0, \omega_0) = ((0, 0, 0), (1, 0, 0))$ и $\sigma(\mathbf{r}) = \sigma \equiv 1$. Введем функционал $I_1 = (F, \tilde{\Phi})$, где $\tilde{\Phi}$ - решение системы (3.2) при $H^T = (1, 0, 0, 0)$, а $F = (\delta(x - x_0), 0, 0, 0)$. Данный функционал I_1 определяет среднее количество столкновений частицы до момента обрыва траектории от источника единичной мощности и, для бесконечной однородной среды, его точное значение известно: $I_1 = 1/\sigma_c$.

В таблице 1 представлены результаты расчетов функционала I_1 для молекулярного рассеяния (3.6) с использованием векторной оценки ζ с ветвлением траектории (см. раздел 2, (3.7)) и переходной плотности

$$p(x, y) = (1 - \sigma_c) e^{-l} r_{11}(\mu) \delta(\omega' - \omega'(\omega, \mu, \varphi)) \delta(\mathbf{r}' - \mathbf{r} - \omega'l)/(2\pi).$$

Использованы следующие обозначения: $E I_1$ - статистическая оценка функционала I_1 , $V I_1$ - статистическая оценка величины $D\zeta$, $\sigma_M = \sqrt{V I_1/M}$ - оценка среднеквадратической погрешности, где M - количество моделируемых “деревьев” траекторий. В каждом варианте расчетов вычислялось среднее количество полного числа столкновений N_1 до обрыва траектории и статистическая оценка среднеквадратической погрешности $\sigma_{N_1} = \sqrt{V N_1/M}$. Каждое “дерево” траекторий моделировалось с использованием “метода поколений” [5], причем размер максимального поколения в процессе расчета при $M = 10^8$ для $I_1 = 12.5$ составил 42, а для $I_1 = 100$ был меньше 281.

Таблица 1. Оценка функционала I_1 при молекулярном рассеянии для различных значений коэффициентов поглощения. Критическое значение $\sigma_c = 0.0787$

M	$\sigma_c = 0.078$ $I_1 = 12.8205$			$\sigma_c = 0.04$ $I_1 = 25$		
	$EI_1 \pm \sigma_M$	VI_1	$EN_1 \pm \sigma_{N_1}$	$EI_1 \pm \sigma_M$	VI_1	$EN_1 \pm \sigma_{N_1}$
10^6	12.8195 ± 0.0181	327	24.9621 ± 0.0393	25.0315 ± 0.0482	2326	53.3291 ± 0.1094
10^7	12.8209 ± 0.0057	328	25.7019 ± 0.0129	25.0287 ± 0.0152	2316	53.3240 ± 0.0345
10^8	12.8215 ± 0.0018	328	25.7034 ± 0.0040	25.0045 ± 0.0048	2315	53.2685 ± 0.0109

M	$\sigma_c = 0.02$ $I_1 = 50$			$\sigma_c = 0.01$ $I_1 = 100$		
	$EI_1 \pm \sigma_M$	VI_1	$EN_1 \pm \sigma_{N_1}$	$EI_1 \pm \sigma_M$	VI_1	$EN_1 \pm \sigma_{N_1}$
10^6	50.1854 ± 0.1356	18399	110.4323 ± 0.3082	100.3426 ± 0.3803	144659	224.4130 ± 0.8644
10^7	49.9485 ± 0.0425	18093	109.8955 ± 0.0966	100.0448 ± 0.1199	143890	223.7361 ± 0.2726
10^8	50.0104 ± 0.0134	18121	100.0051 ± 0.0378	100.0051 ± 0.0378	143443	223.6458 ± 0.0861

Автор выражает благодарность члену-корреспонденту РАН Г.А. Михайлову и д.ф.-м.н. Ухинову С.А. за полезные советы и замечания.

Список литературы

- [1] Марчук Г.И., Михайлов Г.А., Назаралиев М.А. и др. *Метод Монте-Карло в атмосферной оптике.* - Новосибирск: Наука, 1976 [Engl.transl.: Springer-Verlag, 1980]
- [2] Михайлов Г.А. *Оптимизация весовых методов Монте-Карло.* М.: Наука, 1987 [Engl.transl.: Springer-Verlag, 1992].
- [3] Михайлов Г.А., Медведев И.Н. *Оптимизация весовых алгоритмов статистического моделирования.* - Новосибирск: Омега Принт, 2011. - 304 с.
- [4] Г. А. Михайлов, С. А. Ухинов, А. С. Чимаева Дисперсия стандартной векторной оценки метода Монте-Карло в теории переноса поляризованного излучения// Ж. вычисл. матем. и матем. физ., Т.46, № 11, 2006, 2099–2113
- [5] Ермаков С.М., Михайлов Г.А. Статистическое моделирование. М, Наука, 1982.
- [6] I.N.Medvedev, G.A.Mikhailov. Probabilistic-algebraic algorithms of Monte Carlo methods // Russian Journal of Numerical Analysis and Mathematical Modelling. - 2011. - Vol. 26, № 3. - P. 323-336
- [7] I.N. Medvedev, Vector estimators of the Monte Carlo method with a finite variance //Russian Journal of Numerical Analysis and Mathematical Modelling, 2013, Vol. 28, No. 3, P. 231–244.