УДК 519.245 Вестник СПбГУ. Математика. Механика. Астрономия. 2021. Т. 8 (66). Вып. 1
MSC 65C05

О методе Монте-Карло для решения больших систем линейных обыкновенных дифференциальных уравнений

С. М. Ермаков, М. Г. Смиловицкий

Санкт-Петербургский государственный университет,
Российская Федерация, 199034, Санкт-Петербург, Университетская наб., 7–9

Для цитирования: *Ермаков С. М., Смиловицкий М. Г.* О методе Монте-Карло для решения больших систем линейных обыкновенных дифференциальных уравнений // Вестник Санкт-Петербургского университета. Математика. Механика. Астрономия. 2021. Т. 8 (66). Вып. 1. С. 37–48. https://doi.org/10.21638/spbu01.2021.104

В статье рассматривается применение метода Монте-Карло к решению задачи Коши для больших систем линейных дифференциальных уравнений. В первой части статьи дается краткий обзор уже известных результатов применения метода для решения интегральных уравнений Фредгольма. В основной части статьи разбирается применение подхода к системе линейных ОДУ, которая приводится к эквивиалентной системе интегральных уравнений Вольтерра. Это позволяет снять ограничения, связанные со сходимостью мажорантного процесса. Формулируются следующие ключевые теоремы. Теорема 1 указывает требуемые условия согласования, которым должны отвечать переходная и начальная плотности распределения, инициирующие соответствующую цепь Маркова, для которой выполняется равенство между математическим ожиданием оценки и интересующим нас функционалом. Теорема 2 формулирует выражение для дисперсии оценки, в то время как теорема 3 указывает параметры цепи Маркова, минимизирующие значение дисперсии для оценки функционала. В работе приводятся доказательства всех трех теорем. В практической части предложенный метод применяется к системе линейных ОДУ, описывающих замкнутую систему массового обслуживания из десяти условных машин и семи условных рабочих. Решение приводится как для системы с постоянной матрицей коэффициентов, так и для системы с переменной матрицей, где в зависимости от времени меняется интенсивность выхода машин из строя. Также произведено сравнение решения методом Монте-Карло с решением методом Рунге — Кутта. Все результаты отражены в таблицах.

Ключевые слова: метод Монте-Карло, системы ОДУ, интегральное уравнение, задачи массового обслуживания, оптимальная плотность, несмещенная оценка, статистическое моделирование.

1. Введение. Метод Монте-Карло для интегральных уравнений и для больших систем линейных алгебраических уравнений достаточно подробно описан в литературе, например [1, 2], а также [3] и [4]. Иным образом дело обстоит с задачами Коши для больших систем обыкновенных дифференциальных уравнений. Вместе с тем даже некоторые классы линейных систем представляют собой значительный интерес для приложения. В качестве примера можно привести задачи массового обслуживания, которые могут описываться системами с очень большим, или даже бесконечным, числом уравнений. В простейшем случае пуассоновских систем легко получить результат при $t \to \infty$, а для переходного режима записать решение в аналитической форме. Для многих задач удобно использовать прямое имитационное моделирование. Во всех случаях, однако, есть задачи, где перечисленные методы очень трудоемки и практически не применимы. Так, для переходного режима уже в простейшем случае решение выражается бесконечными рядами значений бесселевых функций, а имитационное моделирование в задачах вычисления малых вероятностей может оказаться непомерно трудоемким.

Этих соображений достаточно чтобы рассмотреть отличные от имитационных методы Монте-Карло для решения системы

$$\begin{gathered}\frac{d}{dt}Y(t) = A(t)Y(t) + F(t),\ Y(0) = Y_0,\\ A(t) = ||a_{ij}(t)||_{i,j=1}^{n},\ Y(t) = ||y_1(t), ..., y_n(t)||^T,\ F(t) = ||f_1(t), ..., f_n(t)||^T,\end{gathered} \tag{1}$$

для значений t из промежутка $[0, T]$.

Некоторые подходы к построению алгоритмов методом Монте-Карло обсуждались ранее в работах [5] и [6]. В [5] решались стохастические дифференциальные уравнения, а в [6] рассматривался общий случай нелинейной задачи.

В данной работе рассматривается линейный случай, но с общих позиций, и достаточно полно исследуются вопросы минимизации дисперсии простейшей оценки. Приводятся численные примеры из области теории систем массового обслуживания.

Можно указать много способов приведения задачи (1) к интегральной форме, а методы Монте-Карло решения интегральных уравнений хорошо развиты. По этой причине мы напомним кратко простейшие факты относительно метода Монте-Карло для решения интегральных уравнений 2-го рода.

2. Решение интегральных уравнений Фредгольма методом Монте-Карло. Пусть задано уравнение относительно $\varphi(x)$

$$\varphi(x) = \int k(x, y)\varphi(y)\mu(dy) + f(x), \tag{2}$$

выполняющееся на носителе $\boldsymbol{X}$ вероятностной меры μ, и для множества $\boldsymbol{H}$ функций $h(x)$ таких, что существует интеграл $(\varphi, h) = \int \varphi(x)h(x)\mu(dx)$. Пусть сходится метод последовательных приближений (имеется в виду сходимость по норме в некотором нормированном пространстве, см. [2, с. 272–275]):

$$\tilde{\varphi}_{m+1}(x) = \int |k(x, y)|\tilde{\varphi}_m(y)\mu(dy) + |f(x)|, \quad \tilde{\varphi}_0 = |f(x)|; \quad m = 0, 1, \ldots. \tag{3}$$

Тогда справедливо представление

$$(\varphi, h) = (f, h) + \sum_{l=1}^{\infty} \int \mu_0 ... \int \mu_l h_0, k_{0,1} ... k_{l-1,l} f_l, \tag{4}$$

где обозначаем $\mu_r = \mu(dx_r), h_r = h(x_r), f_r = f(x_r), k_{r-1,r} = k(x_{r-1}, x_r), r = 0, 1, ..., l$.

Выбираем цепь Маркова с начальной плотностью $p^0(x)$ и переходной плотностью $p(x, y)$, удовлетворяющей условию

$$\int p(x, y)\mu(dy) = 1 - g(x), \quad 0 \le g(x) < 1, \tag{5}$$

и условию согласования

$$\begin{aligned} &p^0(x) > 0, \text{ если } h(x) \neq 0,\\ &g(x) > 0, \text{ если } f(x) \neq 0,\\ &p(x,y) > 0, \text{ если } k(x,y) \neq 0. \end{aligned} \tag{6}$$

Здесь и далее подразумеваем $x, y \in \boldsymbol{X}$.

Моделируем цепь Маркова, и на ее траекториях $x_0 \to x_1 \to ... \to x_\tau$ вычисляем функцию

$$J_\tau(x_0, ..., x_\tau) = \frac{h(x_0)k(x_0, x_1)...k(x_{\tau-1}x_\tau)f(x_\tau)}{p^0(x_0)p(x_0, x_1)...p(x_{\tau-1}x_\tau)g(x_\tau)}. \tag{7}$$

Предполагается, что почти все траектории конечные. Это будет так, во всяком случае, если $g(x) > 0$.

Легко показать (см. [1, с. 94–102]), что при выполнении условий сходимости (3) и условий согласования (6) имеет место равенство

$$\mathbf{E}J_\tau(x_0, ..., x_\tau) = (h, \varphi). \tag{8}$$

Для этого достаточно учесть, что плотность вероятности случайной траектории есть $p^0(x_0)p(x_0, x_1), ..., p(x_{\tau-1}, x_\tau)g(x_\tau)$, и воспользоваться определением математического ожидания.

Существует ряд других оценок на траекториях процесса (5), но оценка (7) — одна из простейших в вычислительном отношении оценок, и для нее очень просто получить аналитическое выражение для дисперсии и задать цепь Маркова, удовлетворяющую (3) и (6), для которой значение дисперсии имеет минимальное значение.

Эта оптимальная цепь Маркова является однородной и задается

$$\begin{aligned} &\text{начальной плотностью } \tilde{p}^0(x) = \frac{|h(x)| \cdot \overline{\varphi}(x)}{(|h|, |\overline{\varphi}|)} \text{ и}\\ &\text{переходной плотностью } \tilde{p}(x,y) = \frac{|k(x,y)| \cdot \overline{\varphi}(y)}{\overline{\varphi}(x)}, \end{aligned} \tag{9}$$

где $\overline{\varphi}$ есть предел (3) при $m \to \infty$ в (3).

Интересно отметить (это важно для дальнейшего), что в случае неотрицательных h, f и k оптимальное значение дисперсии оценки (8) оказывается нулем. И хотя выражения (9), определяющие оптимальную цепь Маркова, зависят от точного значения $\overline{\varphi}$, они позволяют сделать некоторые выводы относительно выбора цепи Маркова при решении реальных задач. Так, мы видим что в состав плотности $p(x, y)$ следует включать сомножителем $|k(x, y)|$.

3. Решение задачи Коши для системы обыкновенных дифференциальных уравнений методом Монте-Карло. Вернемся к уравнению (1). Элементы матрицы $A(t)$ и вектора $\tilde{F}(t)$ будем предполагать ограниченными функциями t на вещественной оси. Отсюда следует выполнение условий Липшица для правой части (1) и, следовательно, существование и единственность решения на любом конечном промежутке $[0, T]$. Тогда можем переписать (1) в форме

$$Y(t) = Y(0) + \int_0^t F(\tau)d\tau + \int_0^t A(\tau)Y(\tau)d\tau \tag{10}$$

или

$$Y(t)=\int_0^T A(\theta)E(t-\theta)Y(\theta)d\theta+\tilde{F}(t), \tag{11}$$

где $\tilde{F}(t)=Y_0+\int_0^t F(\tau)d\tau$, а $E(x)=\begin{cases}1, \text{ если } x\geq 0,\\ 0, \text{ если } x<0,\end{cases}$ $\quad t\in[0,T]$.

По аналогии, мы можем выбрать цепь Маркова, удовлетворяющую условиям согласования

$$\begin{aligned}&p^0(i,t)>0, \text{ если } h_i(t)\neq 0,\\ &g(i,t)>0, \text{ если } f_i(t)\neq 0,\\ &p(i,t;j,\theta)>0, \text{ если } a_{i,j}(\theta)\neq 0,\end{aligned} \tag{12}$$

и построить аналог оценки (7) функционала $\int_0^T H(\tau)Y(\tau)d\tau$ для некоторого фиксированного вектора $H(t)=(h_1(t),...,h_n(t))$. Но далее мы замечаем, что уравнение (10) является уравнением вольтерровского типа, и мажорантный процесс

$$\overline{Y}_{m+1}(t)=\int_0^T |A(\theta)|E(t-\theta)\overline{Y}_m(\theta)d\theta+|\tilde{F}(t)| \tag{13}$$

всегда сходится. Это, собственно, следует из теоремы Пикара, но может быть проверено непосредственно.

Определим скалярное произведение двух n-мерных векторов как $(A,B)=a_1b_1+ ...+a_nb_n$. Для построения несмещенной оценки (H,Y) единственным требованием к цепи Маркова остается выполнение условий согласования (12). Если эти условия выполнены, то алгоритм решения задачи состоит в моделировании траектории цепи

$$i_0,t_0\to i_1,t_1\to ...\to i_\tau,t_\tau$$

и вычислении оценки

$$\hat{J}_\tau=\frac{h_{i_0}(t_0)a_{i_0,i_1}(t_1)...a_{i_{\tau-1},i_\tau}(t_\tau)\tilde{f}_{i_\tau}(t_\tau)}{p^0_{i_0}(t_0)p(i_0,t_0;i_1,t_1)...p(i_{\tau-1},t_{\tau-1};i_\tau,t_\tau)g(i_\tau,t_\tau)}, \tag{14}$$

$t_0=T$, для которой выполняется равенство

$$\mathbf{E}\hat{J}_\tau=(H,Y).$$

Оценка $\hat{J}_\tau$ равна нулю вне симплекса $t_0\geq t_1\geq ...\geq t_\tau$. Траектория разыгрывается внутри этого симплекса.

Эти предварительные рассмотрения приводят нас к следующим результатам:

Теорема 1. *Пусть задан случайный процесс Маркова с начальной и переходной плотностями*

$$\begin{aligned}&\boldsymbol{p}^0(t_0)=\overline{p}^0_i(t_0):\{p^0_i(\theta_0)=p^0(i_0,\theta_0)\};\quad i_0=\overline{1,n};\ t_0=T,\\ &\boldsymbol{p}(t)=||p_{i,j}(t)||^n_{i,j=1}:\{p_{i,j}(\theta)=p(i_k,\theta_k;i_{k+1},\theta_{k+1})\};\quad k=\overline{0,\tau};\ t\in[0,T]\end{aligned} \tag{15}$$

соответственно, для которых выполнено условие согласования (12) *и равенство*

$$\sum_{j=1}^n\int_0^t p(i,t;j,\theta)d\theta=1-g(i,t),\quad 0\leq g(i,t)\leq 1. \tag{16}$$

Тогда для оценки (14) *выполнено равенство*

$$\mathbf{E}\hat{J}_\tau = (H, Y)\,, \tag{17}$$

где $H(t)$ *— заданный вектор, а* $Y(t)$ *— решение уравнения* (1).

Доказательство. В соответствии с теоремой Пикара [7, с. 53] имеем

$$(H, Y)(t) = (H, \tilde{F})(t) + \sum_{\tau=1}^{\infty} \int_0^t d\theta_1 ... \int_0^{\theta_{\tau-1}} d\theta_\tau H(\theta_1) A(\theta_2) ... A(\theta_\tau) \tilde{F}(\theta_\tau) =$$
$$= \sum_{\tau=0}^{\infty} \sum_{i_0=0}^{n} ... \sum_{i_\tau=0}^{n} \int_0^t d\theta_0 ... \int_0^{\theta_{\tau-1}} d\theta_\tau \, h_{i_0}(\theta_0) a_{i_0,i_1}(\theta_1) ... a_{i_{\tau-1},i_\tau}(\theta_\tau) \tilde{f}_{i_\tau}(\theta_\tau).$$

Имея заданную цепь Маркова, получаем N траекторий $i_0, t_0 \to i_1, t_1 \to ... \to i_\tau, t_\tau$ с плотностью

$$p^0(i_0, t_0) p(i_0, t_0; i_1, t_1) ... p(i_{\tau-1}, t_{\tau-1}; i_\tau, t_\tau) g(i_\tau, t_\tau). \tag{18}$$

Совокупность этих плотностей определяет вероятностную меру $P(d\theta)$ на множестве траекторий. По условию (16) почти все траектории являются конечными, поэтому выполняется равенство

$$\sum_{\tau=0}^{\infty} \sum_{i_0=0}^{n} ... \sum_{i_\tau=0}^{n} \int_0^t d\theta_0 ... \int_0^{\theta_{\tau-1}} d\theta_\tau p^0(i_0, \theta_0) p(i_0, t_0; i_1, \theta_1) ... p(i_{\tau-1}, t_{\tau-1}; i_\tau, \theta_\tau) g(i_\tau, t_\tau) = 1.$$

Пользуясь определением математического ожидания, при выполнении условий согласования (12) для оценки (14) получаем

$$\sum_{\tau=0}^{\infty} \sum_{i_0=0}^{n} ... \sum_{i_\tau=0}^{n} \int_0^t d\theta_0 ... \int_0^{\theta_{\tau-1}} d\theta_\tau \hat{J}_\tau p^0(i_0, \theta_0) p(i_0, t_0; i_1, \theta_1) ... p(i_{\tau-1} t_{\tau-1}; i_\tau, \theta_\tau) g(i_\tau, t_\tau) =$$
$$= \sum_{\tau=0}^{\infty} \sum_{i_0=0}^{n} ... \sum_{i_\tau=0}^{n} \int_0^t d\theta_0 ... \int_0^{\theta_{\tau-1}} d\theta_\tau h_{i_0}(\theta_0) a_{i_0,i_1}(\theta_1) ... a_{i_{\tau-1},i_\tau}(\theta_\tau) \tilde{f}_{i_\tau}(\theta_\tau).$$

□

Очевидным отличием от рассматриваемого в начале статьи классического случая является:

1) выполнение условия (16);

2) автоматическое выполнение мажорантного условия (13).

При выполнении условий теоремы 1 справедлива также

Теорема 2. *Дисперсия оценки* (14) *конечна и равна*

$$\left(\frac{H_i^2}{\boldsymbol{p}_i^0}, \Xi \right) - (H, Y)^2, \tag{19}$$

где $\Xi(t)$ *есть решение уравнения* $\Xi(t) = \int_0^T \dfrac{A_{ij}^2(\theta) E^2(t-\theta)}{\boldsymbol{p}_{ij}(\theta)} \Xi(\theta) d\theta + \dfrac{\tilde{F}_i^2(t)}{\boldsymbol{g}_i(t)}$, *а нотация* A_{ij}^2; $\tilde{F}_i$; $\boldsymbol{p}_i^0$; $\boldsymbol{p}_{ij}$; $\boldsymbol{g}_i$ *подразумевает покомпонентные операции над матрицами и векторами (операции Адамара).*

Доказательство. Воспользуемся стандартным выражением для дисперсии:

$$D\hat{J}_\tau = \mathbf{E}\hat{J}_\tau^2 - [\mathbf{E}\hat{J}_\tau]^2.$$

Из результатов теоремы 1 мы знаем, что $[\mathbf{E}\hat{J}_\tau]^2 = (H, Y)^2$. Обратимся к выражению $\mathbf{E}\hat{J}_\tau^2$. Возводя $\hat{J}_\tau$ в квадрат, получаем выражение

$$\mathbf{E}\hat{J}_\tau^2(t) = \sum_{\tau=0}^{\infty}\sum_{i_0=0}^{n}\dots\sum_{i_\tau=0}^{n}\int_0^t d\theta_0 \dots$$
$$\dots\int_0^{\theta_{\tau-1}} d\theta_\tau \frac{h_{i_0}^2(\theta_0)a_{i_0,i_1}^2(\theta_1)\dots a_{i_{\tau-1},i_\tau}^2(\theta_\tau)\tilde{f}_{i_\tau}^2(\theta_\tau)}{p^0(i_0,\theta_0)p(i_0,t_0;i_1,\theta_1)\dots p(i_{\tau-1},t_{\tau-1};i_\tau,\theta_\tau)g(i_\tau,\theta_\tau)} \quad (20)$$

для скалярного произведения $\left(\frac{H_i^2}{\boldsymbol{p}_i^0}, \Xi\right)$, где $\Xi(t) = \int_0^T \frac{A_{ij}^2(\theta)E^2(t-\theta)}{\boldsymbol{p}_{ij}(\theta)}\Xi(\theta)d\theta + \frac{\tilde{F}_i^2(t)}{\boldsymbol{g}_i(t)}$, или равносильную систему дифференциальных уравнений

$$\frac{d}{dt}\Xi(t) = \frac{A_{ij}^2}{\boldsymbol{p}_{ij}}(t)\Xi(t) + \frac{F_i^2}{\boldsymbol{g}_i}(t), \quad \Xi(0) = \Xi_0.$$

В соответствии с теоремой Пикара, решение этой системы уравнений существует и единственно:

$$\left(\frac{H_i^2}{\boldsymbol{p}_i^0}, \Xi\right)(t) = \sum_{\tau=1}^{\infty}\int_0^t d\theta_0 \dots \int_0^{\theta_{\tau-1}} d\theta_\tau \frac{H_i^2(\theta_0)}{\boldsymbol{p}_i^0(\theta_0)}\frac{A_{ij}^2(\theta_1)}{\boldsymbol{p}_{ij}(\theta_1)}\dots\frac{A_{ij}^2(\theta_\tau)}{\boldsymbol{p}_{ij}(\theta_\tau)}\frac{\tilde{F}_i^{\,2}(\theta_\tau)}{\boldsymbol{g}_i(\theta_\tau)}.$$

Таким образом, $\left(\frac{H_i^2}{\boldsymbol{p}_i^0}, \Xi\right) = \mathbf{E}\hat{J}_\tau^2$. □

Очевидно, дисперсия оценки всегда конечна, и можно указать такие параметры цепи Маркова, которые обеспечат минимальное значение дисперсии оценки.

Теорема 3. *Дисперсия оценки $\hat{J}_\tau$ принимает минимальное значение при выборе следующих параметров цепи Маркова:*

$$\begin{aligned} q_{onm}^0(i,t) &= \frac{|h_i(t)|\overline{y}_i(t)}{(|H|, \overline{Y})}; \\ q_{onm}(i,t;j,\theta) &= \frac{|a_{i,j}(\theta)|E(t-\theta)\overline{y}_j(\theta)}{\overline{y}_i(t)}, \end{aligned} \quad (21)$$

где $\overline{Y}$ — решение уравнения $\overline{Y}(t) = \int_0^T |A(\theta)|E(t-\theta)\overline{Y}_m(\theta)d\theta + |\tilde{F}(t)|$, а y_i — i-й компонент $\overline{Y}$.

Доказательство. Обратимся к оценке (14). Насколько мы видели, с помощью этой оценки для каждой цепи Маркова, удовлетворяющей условиям согласования, имеем

$$(H, Y)(t) = \mathbf{E}\hat{J}_\tau(t) = \int_0^t \hat{J}_\tau(\omega_{\tau,\theta})P(d\theta),$$

где P — мера на траектории марковской цепи, такая что плотность распределения траектории по отношению к мере $\mu^{\tau+1}$ есть выражение (18). Предположим, что мы хотим ввести другую меру $Q(d\theta)$, такую что производная Радона—Никодима $\dfrac{dP}{dQ}$ существует. Для этого мера P должна быть абсолютно непрерывна по отношению к мере Q. Другими словами, она должна иметь сходную с P структуру: траектория $\omega_{\tau,\theta}$ имеет плотность $q_\tau(i_o, t_0; ...; i_\tau, t_\tau)$ по отношению к $\mu^{\tau+1}$, и q_τ должна быть больше нуля для тех траекторий $\omega_{\tau,\theta}$, для которых выражение плотности (18) принимает строго положительные значения.

В соответствии с теоремой о существенной выборке [1, с. 106], для оптимальной плотности имеем

$$q_{\tau,\text{ опт}}(\omega_{\tau,\theta}) = C|\hat{J}_\tau|p^0(i,t)p(i_0,t_0;i_1,t_1)...p(i_{\tau-1}t_{\tau-1};i_\tau t_\tau),$$

что аналогично

$$q_{\tau,\text{ опт}}(\omega_{\tau,\theta}) = C|h_{i_0}(t_0)a_{i_0,i_1}(t_1)...a_{i_{\tau-1},i_\tau}(t_\tau)\tilde{f}_{i_\tau}(t_\tau)|,$$

где C — константа нормировки,

$$C^{-1} = \sum_{\tau=0}^{\infty}\sum_{i_0=0}^{n}...\sum_{i_\tau=0}^{n}\int_0^t d\theta_0...\int_0^{\theta_{\tau-1}} d\theta_k|h_{i_0}(\theta_0)a_{i_0,i_1}(\theta_1)...a_{i_{\tau-1},i_\tau}(\theta_\tau)\tilde{f}_{i_\tau}(\theta_\tau)|,$$

$$C^{-1} = (|H|, \overline{Y});$$

а $\overline{Y}$ удовлетворяет уравнению (13).

Следует заметить, что необязательно должна существовать цепь Маркова, индуцируемая Q. Однако в рассматриваемом случае представляется возможным указать цепь Макрова, индуцирующую такую оптимальную меру. Из уравнения (13) следует

$$\sum_{j=1}^{n}\int_0^T |a_{i,j}(\theta)|E(t-\theta)\overline{y}_j(\theta)d\theta = \overline{y}_i(t) - |\tilde{f}_i(t)|,\ \frac{|\tilde{f}_i(t)|}{\overline{y}_i(t)} \le 1,$$

$$\sum_{j=1}^{n}\int_0^T \frac{|a_{i,j}(\theta)|E(t-\theta)\overline{y}_j(\theta)}{\overline{y}_i(t)}d\theta = 1 - \frac{|\tilde{f}_i(t)|}{\overline{y}_i(t)}.$$

Принимая $g_{\text{опт}}(i,t) = \dfrac{|\tilde{f}_i(t)|}{\overline{y}_i(t)}$, получаем выражение

$$\sum_{j=1}^{n}\int_0^t q_{\text{опт}}(i,t;j,\theta)d\theta = 1 - g_{\text{опт}}(i,t).$$

Другими словами, условие (16) выполнено для вновь индуцируемой меры Q. Выражение плотности приобретает вид

$$q_{\tau,\text{ опт}}(\omega_\tau) = \frac{|h_{i_0}(t_0)||a_{i_0,i_1}(t_1)|...|a_{i_{\tau-1},i_\tau}(t_\tau)||\tilde{f}_{i_\tau}(t_\tau)|}{(|H|, \overline{Y})}.$$

Из уравнения (13) очевидно, что $q^0_{\text{опт}}(i,t)$ — неотрицательная функция, и она нормирована (то есть является плотностью), и функция $q_{\text{опт}}(i,t;j,\theta)$ неотрицательна. Таким образом, для меры Q выполняются условия согласования (12).

Используя неравенство Коши — Буняковского, находим оценку снизу для $\mathbf{E}\hat{J}_\tau^2$:

$$\int_0^t \left(\hat{J}_\tau \frac{dP}{dQ}(\theta)\right)^2 Q(d\theta) \geq \left(\int_0^t |\hat{J}_\tau \frac{dP}{dQ}(\theta)| Q(d\theta)\right)^2 = (H, Y)^2(t).$$

Подставляя $q_{\tau,\text{ опт}}(\omega_{\tau,\theta})$ в выражение для $\mathbf{E}\hat{J}_\tau^2$, получаем

$$\mathbf{E}\hat{J}_\tau^2(t) = \int_0^t \hat{J}_\tau^2 (\frac{dP}{dQ}(\theta))^2 Q(d\theta) = \int_0^t \hat{J}_\tau^2 \cdot \frac{1}{C^2|\hat{J}_\tau^2|} \cdot C|\hat{J}_\tau|\, p_\tau(\omega_{\tau,\theta}) =$$

$$= \int_0^t \frac{|\hat{J}_\tau|}{C} P(d\theta) = (|H|, \overline{Y})(t) \cdot \int_0^t |\hat{J}_\tau| P(d\theta) = (|H|, \overline{Y})^2(t)$$

и убеждаемся в справедливости теоремы 3. □

Полученные выше результаты позволяют нам вывести

Следствие. *Если $a_{i,j}(t) \geq 0$ и $f_i(t) \geq 0$ для всех i, j и t, то оптимальное значение дисперсии оценки $\hat{J}_\tau$ есть нуль.*

Доказательство. В данном случае очевидно, что при выполнении указанных условий

$$(H, Y)^2 = (|H|, \overline{Y})^2,$$

и тогда в выражении (19) получаем

$$D\hat{J}_\tau = (|H|, \overline{Y})^2 - (H, Y)^2 = 0.$$

□

Заметим, что для справедливости результатов аналогичных теореме 3, зависимость вероятности поглощения от t является существенной. Поясним это на следующем простом примере. Пусть $\hat{F}(t) = 0$, и $A(t) = ||a_{i,j}||_{i,j=1}^n$ не зависит от t. Тогда решение может быть записано в замкнутой форме:

$$Y(t) = \exp(At) = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!}. \tag{22}$$

Скалярное произведение (Y, H) можно вычислить с помощью метода Монте-Карло, моделируя однородную цепь Маркова $\overline{p}^0$, P, для которой выполнены условия согласования (12), и вычисляя оценку

$$\hat{J}_\tau = \frac{h_{i_0}\, a_{i_0,i_1} \ldots a_{i_{\tau-1},i_\tau}\, y^0_{i_\tau}}{p^0_{i_0}\, p_{i_0,i_1} \ldots p_{i_{\tau-1},i_\tau}\, g_{i_\tau} \cdot \tau!}. \tag{23}$$

Легко проверить (см., например, [6]), что при $a_{i,j} \geq 0$, $y^0_{i_\tau} \geq 0$ и $h_{i_0} \geq 0$ нельзя выбрать цепь $\overline{p}^0$; P, обращающую дисперсию в нуль.

Ниже приводятся некоторые примеры решения с помощью предложенного алгоритма уравнений, возникающих в теории массового обслуживания.

4. Постоянная матрица коэффициентов. Приведем иллюстративный пример. Возьмем уравнение системы массового обслуживания, описывающее замкнутую систему из $M = 10$ условных станков, которые выходят из строя с интенсивностью λ, и $S = 7$ рабочих, которые чинят станки с интенсивностью μ. Пользуясь [8, с. 102], получаем следующую систему уравнений:

$$
\begin{aligned}
&\dot{y}_0(t) = -10\lambda\, y_0(t) + \mu/7\, y_1(t),\\
&\dot{y}_1(t) = -[9\lambda + \mu/7]\, y_1(t) + 10\lambda\, y_0(t) + 2\mu/7\, y_2(t),\\
&\dot{y}_2(t) = -[8\lambda + 2\mu/7]\, y_2(t) + 9\lambda\, y_1(t) + 3\mu/7\, y_3(t),\\
&\dot{y}_3(t) = -[7\lambda + 3\mu/7]\, y_3(t) + 8\lambda\, y_2(t) + 4\mu/7\, y_4(t),\\
&\dot{y}_4(t) = -[6\lambda + 4\mu/7]\, y_4(t) + 7\lambda\, y_3(t) + 5\mu/7\, y_5(t),\\
&\dot{y}_5(t) = -[5\lambda + 5\mu/7]\, y_5(t) + 6\lambda\, y_4(t) + 6\mu/7\, y_6(t),\\
&\dot{y}_6(t) = -[4\lambda + 6\mu/7]\, y_6(t) + 5\lambda\, y_5(t) + \mu\, y_7(t),\\
&\dot{y}_7(t) = -[3\lambda + \mu]\, y_7(t) + 4\lambda\, y_6(t) + \mu\, y_8(t),\\
&\dot{y}_8(t) = -[2\lambda + \mu]\, y_8(t) + 3\lambda\, y_7(t) + \mu\, y_9(t),\\
&\dot{y}_9(t) = -[\lambda + \mu]\, y_9(t) + 2\lambda\, y_8(t) + \mu\, y_{10}(t),\\
&\dot{y}_{10}(t) = -\mu\, y_{10}(t) + \lambda\, y_9(t).
\end{aligned}
\tag{24}
$$

В рамках данной статьи был рассмотрен конкретный пример со следующими параметрами: $\mu = 0.0202$ и $\lambda = 0.02$, а также $\mu = 0.0202$ и $\lambda = 0.01$. Такая комбинация параметров дает нам значения коэффициента зарузки (см. [8, с. 104]) $\Psi = 0.99$ и $\Psi = 0.49$ соответственно. При коэффициенте загрузки $\Psi = 0.99$ имитационные методы работают плохо. Второй вариант значения коэффициента загрузки демонстрирует работоспособность предлагаемого метода в сравнении с методом Рунге—Кутта в более конвенциональных условиях. Выберем начальную позицию $y_2^0 = 1$; $y_i^0 = 0$, $\forall i \neq 2$; $T = 1.5$; количество траекторий установим равным $N = 1\,800\,000$.

Для сравнения результатов вычислений используем библиотеку SciPy программного языка Python [9]. Воспользуемся рутиной scipy.integrate.ode("dop853"), которая реализует явный алгоритм Дорманда—Принса порядка $8(5,3)$ для метода Рунге—Кутта. Данный алгоритм является адаптивным алгоритмом, автоматически изменяющим шаг интегрирования в зависимости от того, попадает ли локальная ошибка в требуемое значение погрешности. Погрешность рассчитывается, исходя из формулы $AbsTol + RelTol * |\hat{y}_{step}|$, где $AbsTol$ — абсолютная погрешность, $RelTol$ — относительная погрешность, а $\hat{y}_{step}$ — рассчитанное решение на каждом шаге (см. [10, с. 188–204]). Нами были установлены значения $AbsTol =$ 1e-06; $RelTol =$ 1e-04.

Для проверки точности метода Монте-Карло построим 99.7%-й доверительный интервал $[-\hat{s}_n, \hat{s}_n]$ для оценки $\hat{J}_\tau$ и проверим, попадает ли в него разница между полученными значениями решения. Для иллюстрации внесем в таблицу значения $\hat{s}_n = \dfrac{3\sigma}{\sqrt{N}}$ и значения абсолютной фактической ошибки $\chi = |\hat{y}(t) - y^{RK}(t)|$.

Результаты вычислений приведены в табл. 1 и 2, а также в табл. 3 и 4. При построении таблиц было принято решение не округлять результаты расчетов, полученных методом Рунге—Кутта, но исключить из расчета χ те $y^{RK}(t)$, чья значащая цифра по порядку меньше либо равна значению $AbsTol$. На этих местах в таблице χ стоит прочерк. Нетрудно также увидеть, что значения, полученные методом Монте-Карло, в некоторых состояниях меньше или равны $\hat{s}_n$, что объясняется их близостью к нулю.

Таблица 1. **Сравнение решения методом Монте-Карло и методом Рунге — Кутта при** $T = 1.5$, $\Psi = 0.99$, $N = 1\,800\,000$

	Монте-Карло	Рунге — Кутта
y_0	6.584041e-06	1.425017e-05
y_1	6.602467e-03	6.672860e-03
y_2	7.843779e-01	7.819781e-01
y_3	1.900002e-01	1.898743e-01
y_4	2.019594e-02	2.018617e-02
y_5	1.276377e-03	1.226557e-03
y_6	4.864802e-05	4.658389e-05
y_7	3.336787e-07	1.132339e-06
y_8	1.867346e-08	1.721394e-08
y_9	9.384152e-11	1.495906e-10
y_{10}	$\leq \hat{s}_n$	5.691357e-13

Таблица 2. **Сравнение доверительного интервала метода Монте-Карло и абсолютной погрешности между решениями МК и РК при** $T = 1.5$, $\Psi = 0.99$, $N = 1\,800\,000$

	$\hat{s}_n$	χ
y_0	8.884365e-06	7.666127e-06
y_1	3.728903e-04	7.039296e-05
y_2	4.111465e-03	2.399780e-03
y_3	1.706013e-03	1.258374e-04
y_4	3.551892e-04	9.779131e-06
y_5	5.111231e-05	4.981978e-05
y_6	4.754939e-06	2.064131e-06
y_7	1.866752e-07	–
y_8	1.055185e-08	–
y_9	8.037395e-11	–
y_{10}	1.036435e-16	–

Таблица 3. **Сравнение решения методом Монте-Карло и методом Рунге — Кутта при** $T = 1.5$, $\Psi = 0.49$, $N = 1\,800\,000$

	Монте-Карло	Рунге — Кутта
y_0	1.400258e-05	1.630445e-05
y_1	7.620840e-03	7.575809e-03
y_2	8.787064e-01	8.804753e-01
y_3	1.058588e-01	1.061585e-01
y_4	5.537256e-03	5.601951e-03
y_5	1.644043e-04	1.689383e-04
y_6	2.892959e-06	3.184300e-06
y_7	4.260589e-08	3.841370e-08
y_8	$\leq \hat{s}_n$	2.898100e-10
y_9	$\leq \hat{s}_n$	1.249898e-12
y_{10}	$\leq \hat{s}_n$	2.359078e-15

Таблица 4. **Сравнение доверительного интервала метода Монте-Карло и абсолютной погрешности между решениями МК и РК при** $T = 1.5$, $\Psi = 0.49$, $N = 1\,800\,000$

	$\hat{s}_n$	χ
y_0	4.204214e-06	2.301872e-06
y_1	2.519817e-04	4.503130e-05
y_2	3.694861e-03	1.768928e-03
y_3	8.335724e-04	2.996921e-04
y_4	9.236051e-05	6.469456e-05
y_5	6.750112e-06	4.533990e-06
y_6	3.615292e-07	–
y_7	1.237473e-08	–
y_8	2.560555e-11	–
y_9	5.164668e-14	–
y_{10}	4.225728e-24	–

Таблица 5. **Сравнение решения методом Монте-Карло и методом Рунге — Кутта при** $T = 1.5$, $\lambda(t) = \lambda_0/(1+t^2)$, $\lambda_0 = 0.02$, $N = 2\,200\,000$

	Монте-Карло	Рунге — Кутта
y_0	1.571291e-05	1.569973e-05
y_1	7.285691e-03	7.297974e-03
y_2	8.508264e-01	8.486886e-01
y_3	1.342014e-01	1.343154e-01
y_4	9.442907e-03	9.304705e-03
y_5	3.745529e-04	3.683801e-04
y_6	9.152182e-06	9.115712e-06
y_7	1.262589e-07	1.443672e-07
y_8	1.206253e-09	1.430467e-09
y_9	1.160412e-11	8.104518e-12
y_{10}	1.902090e-16	2.009527e-14

Таблица 6. **Сравнение доверительного интервала метода Монте-Карло и абсолютной погрешности между решениями МК и РК при** $T = 1.5$, $\lambda(t) = \lambda_0/(1+t^2)$, $\lambda_0 = 0.02$, $N = 2\,200\,000$

	$\hat{s}_n$	χ
y_0	1.327831e-07	1.318437e-08
y_1	3.981266e-05	1.228340e-05
y_2	3.025420e-03	2.137891e-03
y_3	1.408177e-03	1.139761e-04
y_4	2.221353e-04	1.382022e-04
y_5	1.686217e-05	6.172786e-06
y_6	8.709463e-07	–
y_7	2.493599e-08	–
y_8	5.125696e-10	–
y_9	7.929071e-12	–
y_{10}	8.923702e-17	–

5. Переменная матрица коэффициентов. Продемонстрируем возможность предложенного метода для решения системы уравнений с матрицей коэффициентов, зависимой от времени. Возьмем систему (24) и сделаем параметр λ зависимым от t: $\lambda(t) = \lambda_0/(1+t^2)$. Сохраним при этом начальный параметр $\lambda_0 = 0.02$. Алгоритм вычисления функционала общего вида (H, Y) повторяет алгоритм, описанный выше, для случая с постоянной матрицей коэффициентов, однако с необходимостью учитывать изменения матрицы коэффициентов $A(t)$ с течением времени.

Оценка $\hat{J}_\tau$ примет вид (14) при условии, что мы оставим прежние параметры для переходной и начальной плотностей распределения, индуцирующих цепь Маркова: $\boldsymbol{p}^0 = \overline{p}_i^0$ для начального распределения; $\boldsymbol{p}(\theta) : \{p(i, \theta_k; j, \theta_{k+1}) = \dfrac{p_{i,j}}{t_i}\}$ для плотности перехода. Расчет проводим с параметрами $\mu = 0.0202$, $\lambda_0 = 0.02$, $T = 1.5$, $y_2^0 = 1$; $y_i^0 = 0$, $\forall i \neq 2$, и $N = 2\,200\,000$. Результаты вычислений продемонстрированы в табл. 5 и 6.

Отметим, что авторы не использовали при вычислениях известные методы понижения дисперсии оценок: методы существенной выборки, ветвления траекторий, квадратурные формулы со случайными узлами, — которые очевидным образом могли бы сократить необходимое время вычислений. Однако приводимые результаты уже в достаточной мере свидетельствуют о работоспособности описанной в статье теории.

Литература

1. Ермаков С. М. *Метод Монте-Карло в вычислительной математике.* Санкт-Петербург, Невский Диалект (2009).

2. Ермаков С. М. *Метод Монте-Карло и смежные вопросы.* Москва, Наука (1975).

3. Ермаков С. М., Сипин А. С. *Метод Монте-Карло и параметрическая разделимость алгоритмов.* Санкт-Петербург, Изд-во С.-Петерб. ун-та (2014).

4. Ермаков С. М., Михайлов Г. А. *Статистическое моделирование.* Москва, Наука (1982).

5. Ermakov S., Pogosian A. On solving stochastic differential equations. *Monte Carlo Methods and Applications* **25**, iss. 2, 155–161 (2019). https://doi.org/10.1515/mcma-2019-2038

6. Ермаков С. М., Товстик Т. М. Метод Монте-Карло для решения систем ОДУ. *Вестник Санкт-Петербургского университета. Математика. Механика. Астрономия* **6 (64)**, вып. 3, 411–421 (2019). https://doi.org/10.21638/11701/spbu01.2019.306

7. Бибиков Ю. Н. *Курс обыкновенных дифференциальных уравнений.* Санкт-Петербург, Лань (2011).

8. Кофман А., Крюон Р. *Массовое обслуживание: теория и приложения*, пер. с франц. Москва, Мир (1965).

9. Virtanen P., Gommers R., Oliphant T. E. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020). https://doi.org/10.1038/s41592-019-0686-2

10. Hairer E. *Solving ordinary differential equations I: Nonstiff problems.* Springer (1993).

Статья поступила в редакцию 3 июня 2020 г.;
после доработки 27 июля 2020 г.;
рекомендована в печать 17 сентября 2020 г.

Контактная информация:

Ермаков Сергей Михайлович — д-р физ.-мат. наук, проф.; sergej.ermakov@gmail.com
Смиловицкий Максим Григорьевич — аспирант; smmxgr@gmail.com

Monte-Carlo for solving large linear systems of ordinary differential equations

S. M. Ermakov, M. G. Smilovitskiy

St. Petersburg State University, 7–9, Universitetskaya nab., St. Petersburg, 199034, Russian Federation

For citation: Ermakov S. M., Smilovitskiy M. G. Monte-Carlo for solving large linear systems of ordinary differential equations. *Vestnik of Saint Petersburg University. Mathematics. Mechanics. Astronomy*, 2021, vol. 8 (66), issue 1, pp. 37–48. https://doi.org/10.21638/spbu01.2021.104 (In Russian)

Monte-Carlo approach towards solving Cauchy problem for large systems of linear differential equations is being proposed in this paper. Firstly, a quick overlook of previously obtained results from applying the approach towards Fredholm-type integral equations is being made. In the main part of the paper, a similar method is being applied towards a linear system of ODE. It is transformed into an equivalent system of Volterra-type integral equations, which relaxes certain limitations being present due to necessary conditions for convergence of majorant series. The following theorems are being stated. Theorem 1 provides necessary compliance conditions that need to be imposed upon initial and transition distributions of a required Markov chain, for which an equality between estimate's expectation and a desirable vector product would hold. Theorem 2 formulates an equation that governs estimate's variance, while theorem 3 states a form for Markov chain parameters that minimise the variance. Proofs are given, following the statements. A system of linear ODEs that describe a closed queue made up of ten virtual machines and seven virtual service hubs is then solved using the proposed approach. Solutions are being obtained both for a system with constant coefficients and time-variable coefficients, where breakdown intensity is dependent on t. Comparison is being made between Monte-Carlo and Rungge — Kutta — obtained solutions. The results can be found in corresponding tables.

Keywords: Monte-Carlo, ODE system, integral equation, queuing theory, optimal density, unbiased estimate, statistical modelling.

References

1. Ermakov S. M. *Monte-Carlo for quantitative mathematics.* St Petersburg, Nevsky Dialect Publ. (2009). (In Russian)

2. Ermakov S. M. *Monte-Carlo and related topics.* Moscow, Nauka Publ. (1975). (In Russian)

3. Ermakov S. M., Sipin A. S. *Monte-Carlo and parametrical divisibility of algorithms.* St. Petersburg, St. Petersburg Univ. Press (2014). (In Russian)

4. Ermakov S. M., Mikhailov G. A. *Statistical modelling.* Moscow, Nauka Publ. (1982). (In Russian)

5. Ermakov S., Pogosian A. On solving stochastic differential equations. *Monte Carlo Methods and Applications* **25**, iss. 2, 155–161 (2019). https://doi.org/10.1515/mcma-2019-2038

6. Ermakov S. M., Tovstik T. M. Monte-Carlo method for solution of systems ODE. *Vestnik of Saint Petersburg University. Mathematics. Mechanics. Astronomy* **6 (64)**, iss. 3, 411–421 (2019). https://doi.org/10.21638/11701/spbu01.2019.306 (In Russian) [Engl. transl.: *Vestnik St. Petersb. Univ. Math.* **52**, iss. 3, 272–280 (2019). https://doi.org/10.1134/S1063454119030087].

7. Bibikov J. N. *A course in ordinary differential equations.* St. Petersburg, Lan' Publ. (2011). (In Russian)

8. Kaufmann A., Cruon R. *Los fenomenos de espera; teoria y aplicaciones.* Continental (1966). [Russ. ed.: *Massovoe obsluzhivanie: teorija i prilozhenija.* Moscow, Mir Publ. (1965)].

9. Virtanen P., Gommers R., Oliphant T. E. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020). https://doi.org/10.1038/s41592-019-0686-2

10. Hairer E. *Solving ordinary differential equations I: Nonstiff problems.* Springer (1993).

Received: June 3, 2020
Revised: July 27, 2020
Accepted: September 17, 2020

Authors' information:

Sergey M. Ermakov — sergej.ermakov@gmail.com
Maxim G. Smilovitskiy — smmxgr@gmail.com